

Table of Contents

- Lec 1 Introduction to Nanotechnology and Nanoscale Transport Phenomena. Microscopic Pictures of Heat Carriers
- Lec 2 Characteristic Time and Length, Simple Kinetic Theory, Characteristic
- Lec 3 Schrödinger Equation
- Lec 4 Quantum Wells, Harmonic Oscillators, Rigid Rotors, and Hydrogen Atoms
- Lec 5 Rigid Rotors, Hydrogen Atom, Electronic Levels in One-Dimensional Lattice Chain
- Lec 6 Electronic Energy Levels in Crystals
- Lec 7 Phonon Energy Levels in Crystals, Crystal Structures
- Lec 8 Reciprocal Lattice, X-ray
- Lec 9 Energy Spectrum in Nanostructures, Density of States, Statistical Distributions
- Lec 10 Specific Heat of Molecules, Electrons, Phonons; Blackbody Radiation
- Lec 11 Effects of Nanostructures on Energy Storage, Energy Transfer by Waves, Electron Waves
- Lec 12 Electromagnetic Waves, Reflection of Waves at a Single Interface
- Lec 13 Acoustic Waves, Interference and Tunneling
- Lec 14 Landauer Formalism
- Lec 15 Transport in Carbon Nanotubes (Guest lecture)
- Lec 16 Transition to Particle Description, Liouville Equation
- Lec 17 Boltzmann Equation, Relaxation Time Approximation
- Lec 18 Fourier Law and Newton's Shear Stress Law
- Lec 19 Ohm's Law and Thermoelectric Effect
- Lec 20 Nanostructured Thermoelectrics (Guest lecture)
- Lec 21 Thermoelectric Effect

- Lec 22 Classical Size Effects, Parallel Direction
- Lec 23 Classical Size Effects, Perpendicular Direction
- Lec 24 Liquid, Brownian Motion, Forces and Potentials, Electrokinetics, Surface Tension

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 1

1. Overview for nano sciences
 - 1.1 Length scale
 - 1.2 Examples in microtechnology
 - 1.3 Examples in nanotechnology
 - 1.4 Nano for energy (phonon, electron; wavelength, mean free path)
 - 1.5 Nanoscale heat transfer in devices (e.g., CMOS)
 - 1.6 Nano and microfabrication
 - 1.7 Transport regimes
 - 1.8 Overview of the book chapters and chapters to be covered
2. Classical Laws related to transport
 - 2.1 Heat transfer
 - 2.1.1 Conduction

Fourier's law:

$$q = -k\nabla T \text{ or } q = -k \frac{dT}{dx} \text{ in one dimension}$$

where:

q [W/m²] is heat flux,
 k [W/m-K] is thermal conductivity.

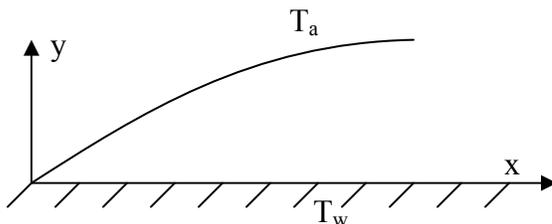
2.1.2 Convection

Newton's law of cooling:

$$q = h(T_w - T_a)$$

where:

h [W/m²K] is heat transfer coefficient.



Nonslip boundary condition is assumed at the wall, i.e.,

$$u_x(y=0) = u_y(y=0) = 0, \quad T(y=0) = T_w.$$

Note: this assumption is NOT accurate for small scales.

2.1.3 Radiation

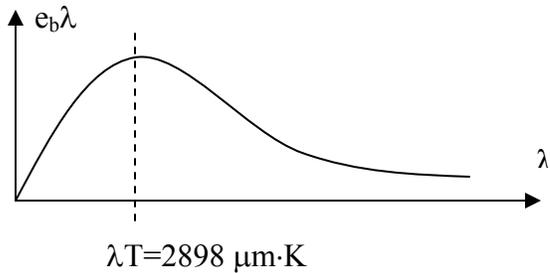
Planck's law:

$$e_{b,\lambda} = \frac{c_1}{\lambda^5 (e^{c_2/\lambda T} - 1)}$$

where:

c_1 and c_2 are constants,
 λ is wavelength, and the subscript b denotes black body.

The curve for block-body radiation is drawn as following:



Integrating the Planck's law leads to the Stefan-Boltzmann law:

$$e_b = \sigma T^4$$

where:

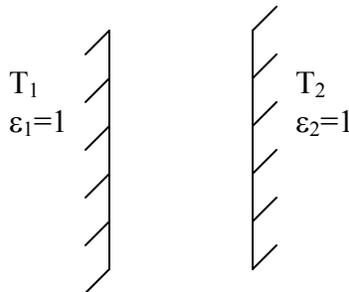
$$\sigma = 5.67 \times 10^{-8} \text{W/m}^2 \text{K}^4.$$

For real surface, we define "emissivity" as

$$\varepsilon = \frac{e}{e_b}.$$

For the two planar walls shown below, the heat flux of radiation is evaluated as

$$q = \sigma(T_1^4 - T_2^4).$$



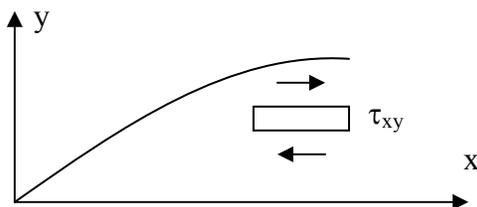
2.2 Newton shear stress law

The shear stress for the sketched one-dimensional flow is:

$$\tau_{xy} = \mu \frac{\partial u}{\partial y}$$

where:

μ [N-s/m²] is dynamic viscosity.



2.3 Fick's diffusion law

$$j_i = -\rho D \frac{dm_i}{dx}$$

where:

D [m^2/s] is mass diffusivity,
 m_i is the mass fraction for the i th species.

2.4 Ohm's law

$$R = \frac{V}{I}$$

or

$$J = \sigma \varepsilon = \sigma \left(-\frac{1}{e} \frac{d\Phi}{dx} \right) = -\sigma \frac{d\varphi_e}{dx}$$

where:

J [A/m^2] is electric current density,
 σ [$\Omega^{-1}\text{m}^{-1}$] is electrical conductivity,
 ε [V/m] is electric field,
 Φ is potential energy,
 φ is electrostatic potential.

2.5 Questions

- What are the similarities among above equations?
- Are these laws still valid at nanoscale?

2.6 Note:

All above are constitutive equations with two unknown variables.
Another equation (e.g. mass, momentum conservation) is needed to solve problems.

3. Scaling trend

For a sphere, the volume-to-surface ratio is

$$\frac{V}{S} = \frac{4\pi r^3 / 3}{4\pi r^2} = \frac{r}{3}$$

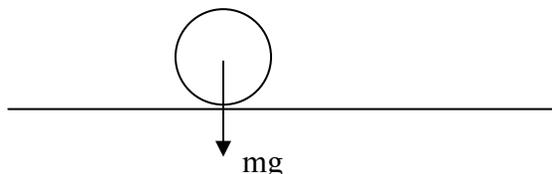
The volumetric effect decreases with the reducing length scale. Surface effect becomes dominant at smaller scales.

For a spherical fluid drop, we have

$$\gamma = \frac{\text{Gravitational force}}{\text{Surface force}} = \frac{\rho g (4\pi r^3 / 3)}{\sigma (2\pi r)} = \frac{2\rho g r^2}{3\sigma}$$

Substituting $\rho = 10^3 \text{ kg}/\text{m}^3$, $\sigma = 78 \text{ mN}/\text{m}$ into this equation, we get

$r = 1 \text{ m}$, $\gamma = 8.4 \times 10^4$; $r = 1 \text{ mm}$, $\gamma = 8.4 \times 10^{-2}$.

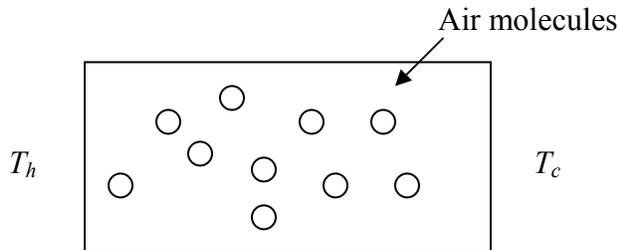


4. Microscopic pictures of energy carriers

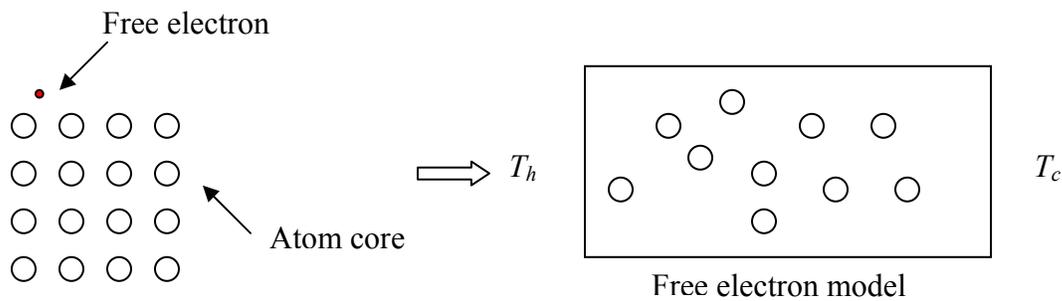
4.1 Heat

4.1.1 Heat conduction

Gases: hotter air molecules (with larger kinetic energy) randomly pass their excess energy to cooler molecules. Heat is transported to the cold side by such a process. Note the average velocity of a molecule can be as large as 500 m/s.



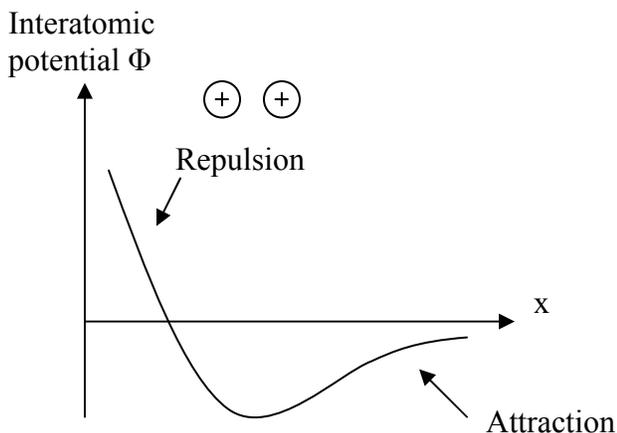
Dielectric solids: heat is conducted through the vibration of atoms. The atom cores are spaced by 2-5 Å in the lattice. Under the free electron approximation, the electrons are viewed as free electron gas.



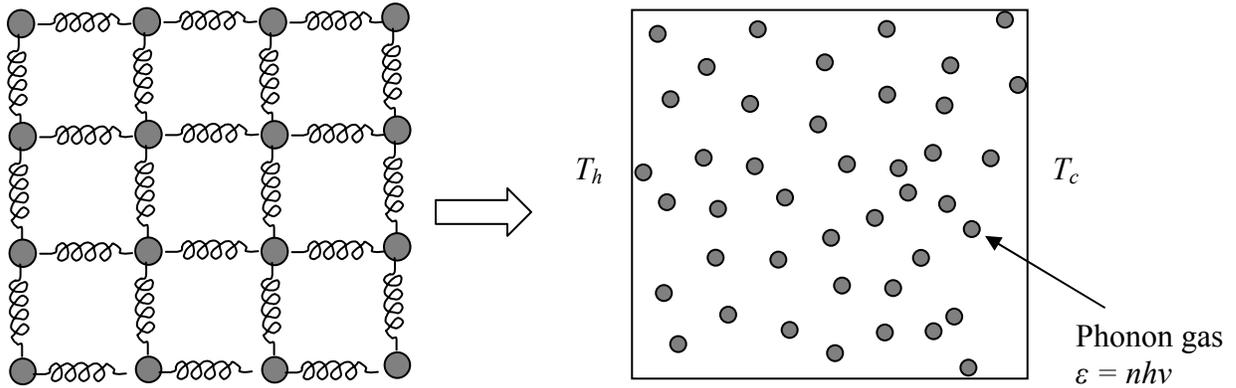
Consider two atoms with a parabolic interatomic potential. The interatomic force is

$$F = -\nabla\Phi \approx K\Delta x$$

where Δx is the displacement from the minimum potential position (equilibrium position), K is constant.



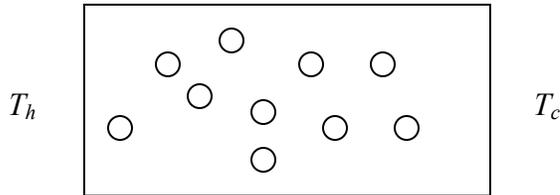
A simplified picture of the interatomic interactions in crystals can be represented by the mass-spring system. The propagation of sound in a solid is due to long wavelength lattice waves. Quantum mechanics states that the energy of each lattice wave is discrete and must be multiples of $h\nu$. Based on argument we will discuss in chapter 5, the spring system can be further simplified as a box of phonon particles.



Now, molecules, electrons, and phonons are all gases in a box. You can see similarities and why I said we can describe them in parallel.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 2

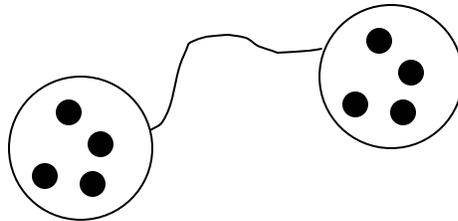
5.1 Heat conduction



In last lecture, we describe electrons as free electron gas and lattice vibrations as phonon gas. Basically they are both gases in a box.

5.2 Convection

- 1) Typically electron velocities are 10^5 - 10^6 m/s, while phonon velocities are the sound velocity.
- 2) In heat conduction processes, the average velocity of heat carriers is $\bar{v} = 0$. In convection, a non-vanishing average velocity superimposed on their random velocity, resulting in $\bar{v} \neq 0$.
- 3) When a liquid or gas molecule is moved from one place to another due to its nonzero velocity, it also carries its internal energy.



5.3 Radiation

- 1) Wavelength comparison

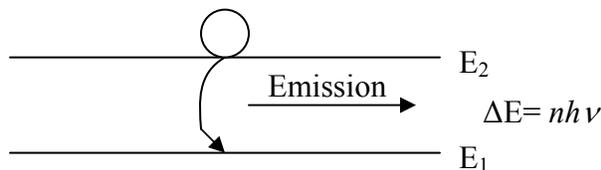
For radio/TV signals, we get

$$\lambda = c / f = \frac{3 \times 10^8}{900 \times 10^6} = \frac{1}{3} \text{ (m)},$$

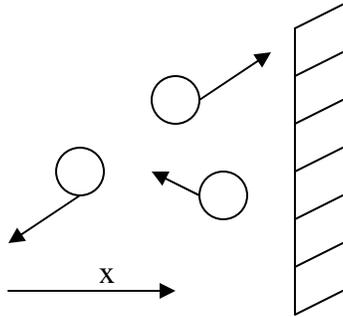
where we use 900 MHz as the frequency. This wavelength is still much larger than that of the thermal radiation (around $0.5 \mu\text{m}$).

- 2) Generation of thermal radiation

Thermal radiation typically refers to the electromagnetic waves that are generated by the oscillation charges in the atoms and crystals, while TV and radio signals are generated by artificial current oscillation in a circuit. An electromagnetic wave at frequency ν can only have energy that is multiple times of $h\nu$.



5.4 Pressure and shear stress



As is shown in the figure, the velocities of gas molecules distribute randomly in all directions. Pressure is caused by their momentum changes normal to the wall. For one molecule, we have

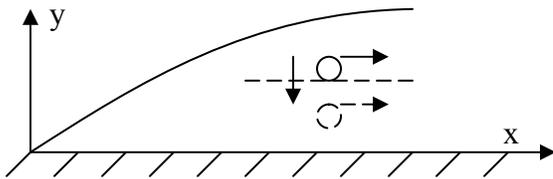
$$\vec{F} = m\vec{a} = \frac{d(m\vec{v})}{dt} \text{ or } F_x = \frac{m(v_{x>0} - v_{x<0})}{\Delta t} = \frac{m\Delta v_x}{\Delta t}.$$

Denote n [m^{-3}] as the number of particles per unit volume. We notice nv_x [$\text{m}^{-2}\text{s}^{-1}$] has the physical meaning as the flux of particles on the wall. Assuming elastic collisions between the wall and molecules, we have $\Delta v_x = 2v_{x>0}$. Thus

$$P = \frac{1}{2} nv_{x>0} (m\Delta v_x) = mnv_{x>0}^2 = mnv_x^2 = mn \frac{v^2}{3},$$

in which we use average $v^2 = v_x^2 + v_y^2 + v_z^2 = 3v_x^2$. From here, you can derive the ideal gas law using the relationship between velocity and temperature that I will talk below.

For assignment 1, similar processes can be followed to calculate the shear stress.



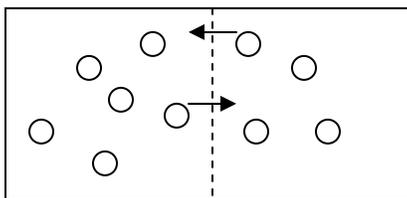
5.5 Charge transport

Similarly, we have

$$\vec{F} = q\vec{e} = -e\vec{e}; \quad \vec{J} = \sigma\vec{e},$$

where \vec{e} is electrical field and \vec{J} is electric current density. An expression of σ can be derived.

5.6 Mass diffusion

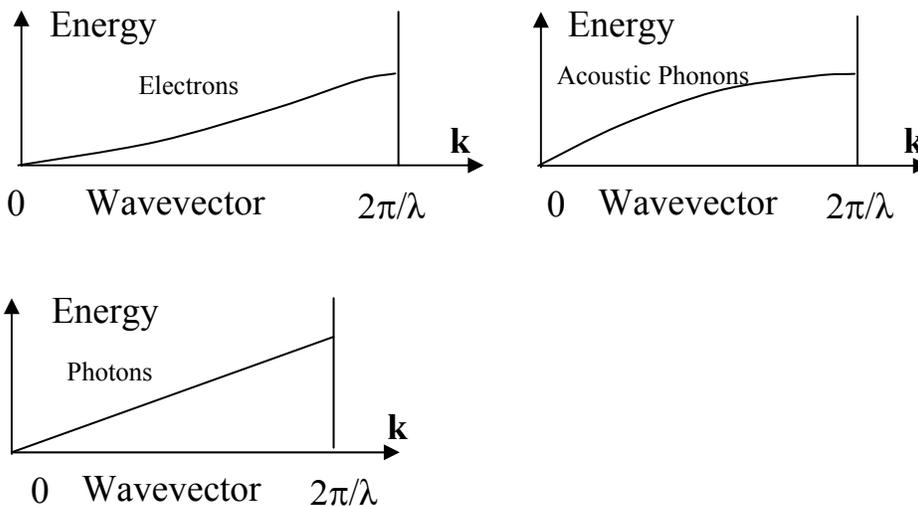


6. To understand transport and energy conversion, we need to know:
- How much energy/momentum can a particle have?
 - How many particles have the specified energy E ?
 - How fast do they move?
 - How far can they travel?
 - How do they interact with each other?

6.1 How much energy/momentum can a particle have?

	Classical mechanics	Quantum mechanics
Energy	$E = E_{\text{Kinetic}} + E_{\text{Potential}}$	E is the eigenvalue of the Schrödinger equation
E_{Kinetic}	$E_{\text{Translation}} = \frac{mv^2}{2}$ (for quantum case, I did not give answer but point out it is the solution for particle in a box)	
	$E_{\text{Vibration}} = \frac{mv^2}{2} + \frac{Kx^2}{2}$	$E_{\text{Vibration}} = h\nu(n + \frac{1}{2}); \nu = \frac{1}{2\pi} \sqrt{\frac{K}{m}}$ ($n=0, 1, 2, \dots$)
	$E_{\text{Rotation}} = \frac{I\omega^2}{2}$	$E_{\text{Rotation}} = hBl(l+1)$ ($l=0, 1, \dots$)

Here the Planck constant $h = 6.6 \times 10^{-34} \text{ J}\cdot\text{s}$. The vibration energy of a standing wave inside the potential well is discrete in quantum mechanics. In the table, we give the allowed energy levels of a harmonic oscillator, which is a model for the vibrations of a diatomic molecule such as H_2 . The dispersion relation (E - \mathbf{k} relation) for electrons, phonons, and photons are sketched in following figures, in which the wavevector \mathbf{k} points to the direction of wave propagation (electron, photon, and phonon waves).



For photons, the energy is just a linear function of the wavevector, i.e.,

$$E = hc / \lambda = \frac{hc}{2\pi} \frac{2\pi}{\lambda} = \frac{hc}{2\pi} |\vec{k}|.$$

6.2 How many particles have the specified energy E?

For a monoatomic ideal gas system, the only energy of each atom is their kinetic energy,

$$E = \frac{m}{2} (v_x^2 + v_y^2 + v_z^2).$$

Statistical thermodynamics gives the probability density $f(E)$, defined as the probability of finding the carriers at energy E per energy interval surrounding E, that a particle in an equilibrium system at a temperature T as,

$$f(E) = Ae^{-E/(k_B T)},$$

where the Boltzmann constant $k_B = 1.38 \times 10^{-23}$ J/K.

6.3 How fast do they move?

First we use normalization of f to calculate A. Since the probability of finding this particle having energy between 0 and infinity must be one, we have

$$\int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} f(v_x, v_y, v_z) dv_z = 1.$$

Using spherical coordinates can simplify the calculation as

$$\int_0^{\infty} 4\pi v^2 f(v) dv = 1.$$

Both of above equations yield

$$A = \left(\frac{m}{2\pi k_B T} \right)^{3/2}.$$

Thus

$$f(\mathbf{v}) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} \exp \left[-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T} \right],$$

which is also called the Maxwell distribution. The average energy of the monoatomic gas is

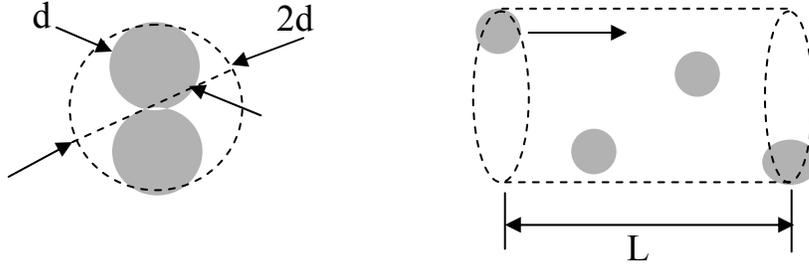
$$\begin{aligned} \langle E \rangle &= \int_{-\infty}^{\infty} dv_x \int_{-\infty}^{\infty} dv_y \int_{-\infty}^{\infty} \frac{m}{2} (v_x^2 + v_y^2 + v_z^2) \left(\frac{m}{2\pi k_B T} \right)^{3/2} \exp \left[-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T} \right] dv_z \\ &= \frac{3}{2} k_B T \end{aligned}$$

At room temperature (300 K), this average energy is 39 meV, or 6.21×10^{-21} J. For He gas, $m = 6.4 \times 10^{-27}$ kg. Using $\frac{mv^2}{2} = \frac{3}{2} k_B T$, we can calculate average $v = 1000$ m/s, while it becomes 500 m/s for air. It is good to remember that $k_B T$ at room temperature is 26 meV.

Note: The flux of quantity X (e.g. momentum, energy) is expressed as $n(X)X \cdot v$, which is already used in 5.4.

6.4 How far can they travel?

The effective diameter for two atoms to collide is $2D$. . If the number concentration is n [m^{-3}], then the number of molecules that this particle will collide with is $n\pi D^2 L$. The average distance L between each collision satisfies $n\pi D^2 L = 1$.



Thus the mean free path is

$$\Lambda = \frac{1}{n\pi D^2}.$$

Noticing $n = \rho/m$, where m is molecular weight, ρ is density. For ideal gases, we have $P = \rho RT$. Therefore,

$$\Lambda = \frac{m}{\pi D^2 \rho} = \frac{mRT}{\pi D^2 P} = \frac{R_u T}{\pi D^2 P N_A} = \frac{k_B T}{\pi D^2 P},$$

where the universal gas constant $R_u = N_A k_B = 8.314 \text{ J/mol} \cdot \text{K}$.

The ideal gas law can also be derived as following:

In section 5.4, we derive $P = mn \frac{v^2}{3}$. In section 6.3, we get $\frac{mv^2}{2} = \frac{3}{2} k_B T$. Thus

$$P = mn \frac{v^2}{3} = nk_B T = \frac{N}{V} k_B T = \frac{N}{N_A} \frac{N_A}{V} k_B T = \frac{N}{N_A} \frac{R_u}{V} T, \text{ or } PV = \bar{N} R_u T, \text{ in which } \bar{N} \text{ is}$$

the mole number.

At room temperature, atmosphere pressure, the mean free path is

$$\Lambda = \frac{k_B T}{\sqrt{2} \pi D^2 P} = \frac{1.38 \text{e-}23 \times 300}{\sqrt{2} \pi \times (2.5 \text{e-}10)^2 \times 1.01 \text{e}5} = 0.14 \mu\text{m},$$

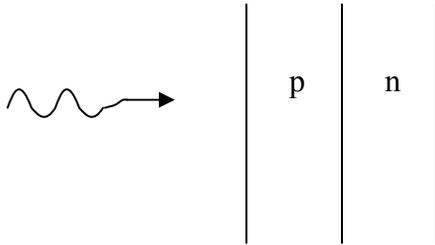
in which $\sqrt{2}$ comes when we consider relative velocity instead of assuming other particles are stationary. The collision time (relaxation time) is $\tau = \Lambda / \bar{v} \sim 10^{-10} \text{ s}$.

For $P=1 \text{ mTorr}$, we have $\Lambda = 0.1 \text{ m}$. The molecules seldom collide with each other.

Note: This relaxation time is not very small for current technology. For short-pulse lasers, the shortest period is only a few fs.

6.5 How do they interact with each other?

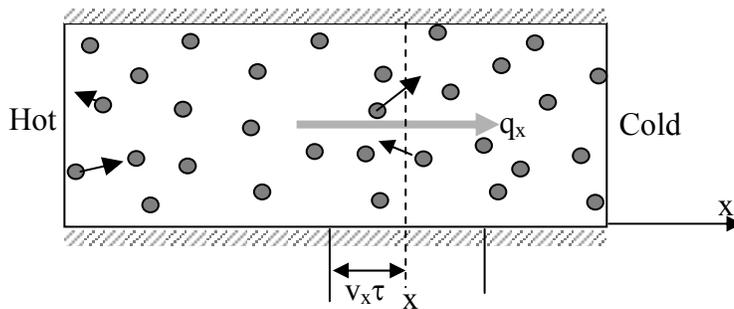
The collisions between particles can be elastic or inelastic. In a solar cell, the photon excites electrons to generate an output voltage. Energy conversion is closely associated with transport.



7. Simple kinetic theory

In the figure, half of the carriers within $v_x \tau$ can go across the interface before being scattered. Here v_x is the x component of the random velocity of the heat carriers and τ is the relaxation time — the average time a heat carrier travels before it is scattered and changes its direction. So the net heat flux carried by heat carriers across the interface is

$$q_x = \frac{1}{2} (nE v_x) \Big|_{x-v_x \tau} - \frac{1}{2} (nE v_x) \Big|_{x+v_x \tau} .$$



Using a Taylor expansion, we can write the above relation as

$$\begin{aligned} q_x &= -v_x \tau \frac{d(nE v_x)}{dx} \\ &= -v_x^2 \tau \frac{d(nE)}{dx} \\ &= -v_x^2 \tau \frac{du}{dT} \frac{dT}{dx} . \end{aligned}$$

Notice specific heat $C = du / dT$, $v_x^2 = v^2 / 3$. The above equation changes into

$$q_x = -\frac{v^2}{3} \tau C \frac{dT}{dx} .$$

Compared with the Fourier's law, we know that thermal conductivity $k = \frac{\tau C v^2}{3}$. The capital C is specific heat per unit volume, $C = \rho c$.

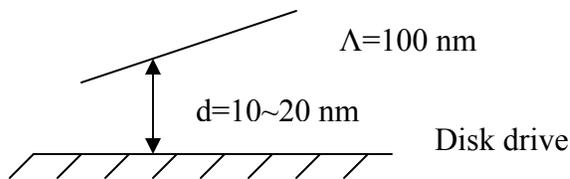
2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 3

8. Micro & Nanoscale Phenomena

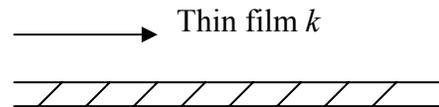
8.1 Classical size effects

In section 7, the characteristic length of the box is much longer than the mean free path Λ . Therefore, the collisions between molecules and the wall are neglected in our derivation and thermal conductivity is regarded as the bulk property of the gas. However, there are many applications in which Λ becomes comparable or larger than the size of the system. The classical size effects occur in such situations.

Example 1: $\Lambda > d$ for a disk drive



Example 2: $\Lambda \sim d$ or $\Lambda > d$ for thin films



In example 2, we can further reduce the film thickness to enhance the size effects. With measured data for k and specific heat c , the mean free path in silicon can be estimated by

$$k = \frac{cv\Lambda}{3},$$

where v is sound velocity. The approximated mean free path Λ is around 40 nm, while the actual value is around 300 nm. The size effects occur for silicon films with thickness less than Λ .

Note: in some thermal insulation applications, we also use porous materials whose pore sizes are comparable to or less than Λ . The thermal conductivity of the air trapped in the pores will be significantly reduced.

8.2 Quantum size effects

According to quantum mechanics, electrons and phonons are also material waves; the finite size of the system can influence the energy transport by altering the wave characteristics, such as forming standing waves and creating new modes that do not exist in bulk materials.

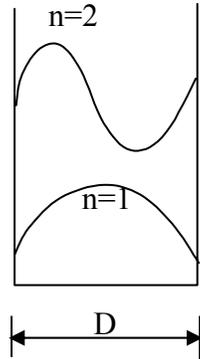
For example, electrons in a thin film can be approximated as standing waves sitting inside a potential well of infinite height. The condition for the formation of such standing waves is that the wavelength, λ , satisfies the following relation

$$D = \lambda n / 2 \quad (n=1,2,\dots),$$

where D is the width of the potential well. According to the de Broglie relation, the wavelength is

$$\lambda = h / p,$$

where h is Planck constant ($h=6.6 \times 10^{-34}$ Js), p is momentum. The energy of the electron is thus $E=p^2/2m$ or $E = \frac{p^2}{2m} = \frac{h^2}{8m} \left(\frac{n}{D}\right)^2$.



1) For a free electron, $D=1$ mm, we have

$$E = \frac{n^2}{8 \times 9.1 \times 10^{-31}} \left(\frac{6.6 \times 10^{-34}}{10^{-3}}\right)^2 \sim 10^{-34} n^2 \ll k_B T = 4.14 \times 10^{-21} J \text{ at room temperature.}$$

2) For $D=1e-8$ m, we calculate $E > k_B T$. Further reducing D results in more observable E.

8.3 Fast transport

For many materials, we have $\tau = 10^{-12} - 10^{-11}$ s. Laser pulse can be as short as a few femtosecond, we cannot use diffusion theory when the time scale is shorter than the relaxation time.

Chapter 2 Material Waves & Energy Quantization

2.1 Basic wave characteristics

2.1.1 Traveling wave

First consider a harmonic wave (such as an electric or a magnetic field) represented by a sine function traveling along the positive x-direction,

$$\vec{\Phi} = A \sin\left(2\pi ft - \frac{2\pi x}{\lambda}\right) \hat{y}$$

$$= A \sin(\omega t - kx),$$

where $\omega = 2\pi f$ denote angular frequency, $k = 2\pi / \lambda$ denote wavevector magnitude, \hat{y} is a unit vector in the y-coordinate direction. Here $2\pi ft$ is the time term, while $\frac{2\pi x}{\lambda}$ is the spatial term.

For constant phase, i.e., $\omega t - kx = \text{const}$, we have

$$\frac{dx}{dt} = \frac{\omega}{k} = f\lambda = v_p,$$

where v_p is phase velocity. It indicates how fast the wave phase is propagating.

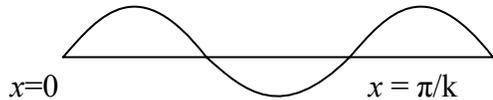
Note: for light we have a linear relationship $\omega = ck$.

2.1.2 Complex representation

It is convenient to use the complex representation of the sine and cosine functions, e.g.

$$A e^{-i(\omega t - kx)} = A [\cos(\omega t - kx) - i \sin(\omega t - kx)].$$

2.1.3 Standing wave



We can create a standing wave by superimposing two traveling waves along the positive and negative x -directions (assuming that the problem is linear such that the superposition principle applies),

$$\Phi = A [\sin(\omega t - kx) - \sin(\omega t + kx)] \hat{y} = -2A \cos(\omega t) \sin(kx) \hat{y},$$

which has fixed nodes in space such that $\Phi=0$. These nodes are similar to the string ends of a guitar.

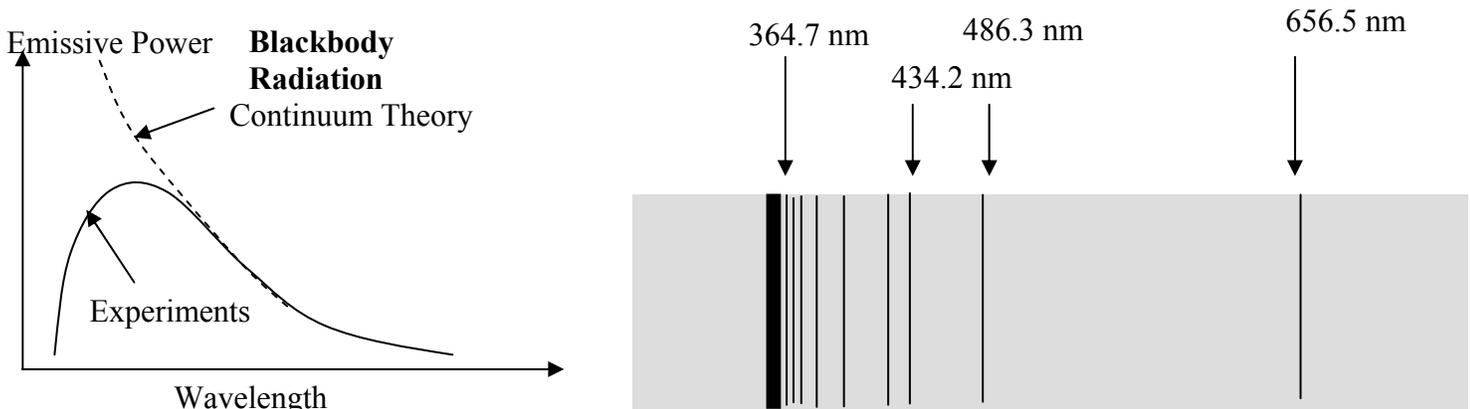
The energy of a wave is typically $U \propto |\Phi|^2$.

2.2 Wave-particle duality

2.2.1 Electromagnetic (EM) wave

Quantum mechanics started with the explanation for blackbody radiation and the absorption spectra of gases. Isaac Newton (1642-1727) believed that radiation was particle-like in nature rather than wave-like, as we are more familiar with today. It was the discovery and explanation of interference and diffraction phenomena, from the work of Christian Huygens (1629-1695), Thomas Young (1773-1829), Augustin Jean Fresnel (1788-1827), and others, followed by Maxwell (1831-1879) and his celebrated equations that solidified the foundation of the wave nature of the electromagnetic field.

The Maxwell equations, however, fail to explain the emission and absorption processes, such as the experimentally observed fine spectra of absorption in various gases, and the blackbody radiation.



To explain the blackbody radiation, Max Planck (1858-1948) introduced a radical hypothesis that the allowable energy of the electromagnetic field at a frequency ν is not continuous, but is a multiple of the following basic energy unit

$$E = nh\nu,$$

in which h is called the Planck constant and has a value $h=6.6 \times 10^{-34}$ J.s.

According to the relativity theory, we have

$$E = mc^2.$$

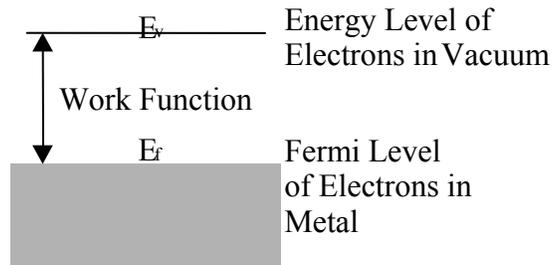
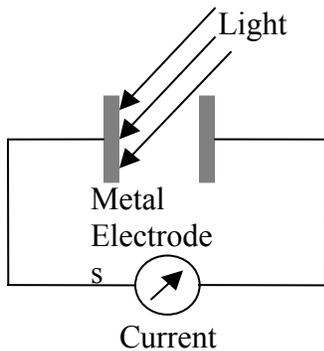
Thus the momentum is

$$p = mc = E/c = h\nu/c = h/\lambda.$$

After 1926, people used the term “photon” to name the quantum with $E = h\nu$, $p = h/\lambda$.

Einstein used the corpuscular characteristics of electromagnetic radiation to explain some puzzling results from the basic photoelectricity experiment in the following figure. Based on the photon particle concept, Einstein reasoned that one photon can excite an electron out of the metal surface only when the photon energy is higher than the work function A ($=E_v - E_f$), which is the energy difference between electrons at the vacuum level, E_v , and inside the metal, E_f , i.e.,

$$h\nu_p \geq E_v - E_f.$$

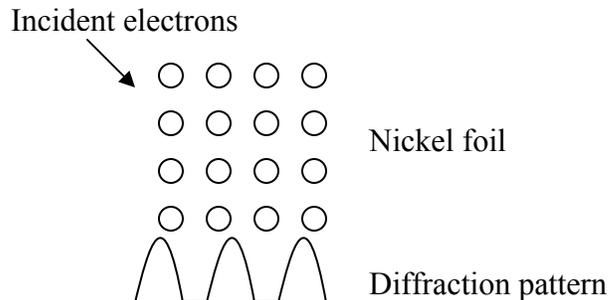


2.2.2 Material wave

The wave-particle duality of light triggered de Broglie to postulate that a material particle also has wave properties. Based on an analogy to the Planck-Einstein relations, he proposed that the wavelength of any particle is

$$\lambda = h/p,$$

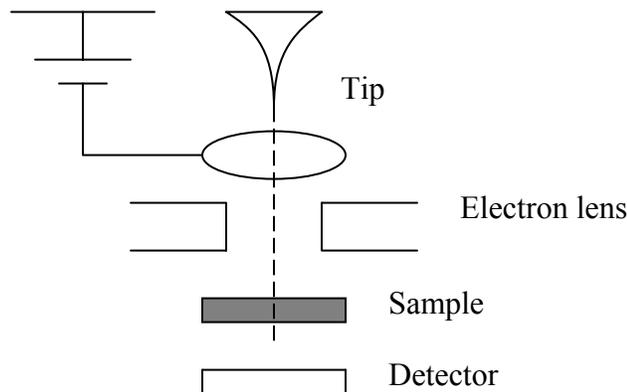
where p is the magnitude of the particle momentum.



An electron with a velocity 1 m/s and a mass of 9.1×10^{-31} kg yields 0.7×10^{-3} m, a quite long wavelength. The first proof of the wave properties of particles came from the electron diffraction experiment, in which peak signals are observed in specified incident angles.

Transmission Electron Microscope (TEM)

A TEM uses thermal excitation or applying a high voltage to draw electrons from the tip end. The electrons are then accelerated by the strong electrical field to gain a large momentum p (small $\lambda \sim 1 \text{ \AA}$). Since the resolution is normally comparable to wavelength λ , high resolution is obtained with electron energy as high as MeV magnitude. The electrons penetrate through the sample (less than 200 nm thick) and the diffraction/transmission is observed from the detector.



Scanning Electron Microscope (SEM)

Different from a TEM, a SEM only observes the surface and electrons do not penetrate the sample. Electrons have lower energy in a SEM.

2.2.3 Mathematical Description of Waves

Two basic methods have been developed to describe the materials waves. The first was the matrix method developed by Heisenberg (1925). Shortly after, Schrödinger developed the famous equation that bears his name. These two descriptions are equivalent among themselves, so we will focus on the Schrödinger equation (Schrödinger, 1926). The equation states that the wavefunction of any matter obeys the equation,

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi_t + U \Psi_t = i\hbar \frac{\partial \Psi_t}{\partial t},$$

where m is the mass, t is the time, U is the potential energy (related to b.c), and Ψ_t is called the wave function of the matter, $\hbar = h/2\pi$. If $U=0$, the Schrödinger equation becomes

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi_t = i\hbar \frac{\partial \Psi_t}{\partial t},$$

which is similar to heat conduction equation $k \nabla^2 T = \rho c \frac{\partial T}{\partial t}$ but the magic imaginary unit “i” really gives rise to wave behavior.

Schrödinger himself did not come up with a correct explanation for the meaning of wavefunction. The right explanation was given by Born, who suggested that Ψ_t itself is not an observable quantity, but $\Psi_t\Psi_t^*$ is the probability density function to find the particle at location x , where “*” means complex conjugate. The normalization requirement for the probability function is then

$$\int \Psi_t^* \Psi_t dV = 1.$$

The expectation value of any quantity (such as energy, momentum, location, etc.) can be calculated from

$$\langle \Omega \rangle = \int \Psi_t^* \Omega \Psi_t dV.$$

In quantum mechanics, quantities such as location, energy, and momentum of matter should be understood in terms of probability values. These quantities are expressed by operators, such as position operator:

$$\Omega = \mathbf{r}$$

time operator:

$$\Omega = t$$

momentum operator:

$$\Omega = \mathbf{p} = -i\hbar\nabla$$

$$= -i\hbar\left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}\right) = p_x\hat{x} + p_y\hat{y} + p_z\hat{z}$$

and the energy operator

$$\begin{aligned} \Omega = H &= \frac{\mathbf{p} \cdot \mathbf{p}}{2m} + U = \frac{\mathbf{p}^2}{2m} + U \\ &= -\frac{\hbar^2}{2m}\nabla^2 + U = -\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} + \frac{\partial}{\partial z^2}\right) + U \end{aligned}$$

Note: in equation $\langle \Omega \rangle = \int \Psi_t^* \Omega \Psi_t dV$, you cannot switch Ψ_t^* and Ψ_t if Ω contains the gradient operator ∇ and the Laplace operator ∇^2 .

Standard deviation:

Similar to

$$\Delta x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \langle x \rangle)^2},$$

in quantum mechanics we have

$$\langle \Delta q \rangle = \sqrt{\int \Psi_t^* (q - \langle q \rangle)^2 \Psi_t dV}.$$

Heisenberg uncertainty principle states

$$\langle \Delta p \rangle \langle \Delta x \rangle \geq \frac{\hbar}{2}; \langle \Delta t \rangle \langle \Delta E \rangle \geq \frac{\hbar}{2}.$$

From the Schrödinger equation, we can derive

$$\frac{\partial |\Psi_t|^2}{\partial t} + \nabla \cdot \vec{J} = 0,$$

where the first term means the changing rate of density, the second flux term is

$$\vec{J} = \frac{i\hbar}{2m} (\Psi_t \nabla \Psi_t^* - \Psi_t^* \nabla \Psi_t).$$

Note: please compare this to continuity equation $\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0$.

Separation of variables:

Assuming $\Psi_t(\mathbf{r}, t) = \Psi(\mathbf{r})Y(t)$ and substituting into the Schrödinger equation, we get

$$\frac{1}{\Psi} \left[-\frac{\hbar^2}{2m} \nabla^2 \Psi + U\Psi \right] = i\hbar \frac{1}{Y} \frac{dY}{dt} = E,$$

where E is a constant (eigenvalue) since Ψ depends on \mathbf{r} only and Y depends on t only.

Its meaning will be explained later ($\langle H \rangle = E$). Solving for Y leads to

$$Y = C_1 \exp\left[-i\frac{E}{\hbar}t\right] = C_1 \exp[-i\omega t].$$

And the governing equation for $\Psi(\mathbf{r})$ is called the steady-state Schrödinger equation

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi + (U - E)\Psi = 0.$$

This is an eigen equation with the eigenvalue E and eigenfunction Ψ determined by the potential profile U and the boundary conditions.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 4

Quick review of Lecture 3

Photon: $E = h\nu$, $p = h / \lambda$.

Assuming $\Psi_t(\mathbf{r},t)=\Psi(\mathbf{r})Y(t)$, we use separation of variables to solve the Schrödinger equation

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi_t + U \Psi_t = i\hbar \frac{\partial \Psi_t}{\partial t}.$$

The solutions are

$$Y = C_1 \exp\left[-i \frac{E}{\hbar} t\right] = C_1 \exp[-i\omega t],$$

and

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi + (U - E) \Psi = 0,$$

where the eigen value E represents the total energy of the system.

Heisenberg uncertainty principle states

$$\langle \Delta p \rangle \langle \Delta x \rangle \geq \frac{\hbar}{2}; \langle \Delta t \rangle \langle \Delta E \rangle \geq \frac{\hbar}{2}.$$

2.3 Example solutions:

Here we determine $u(\vec{r})$ by the boundary conditions and will not consider the $u(\vec{r},t)$ case.

2.3.1 Free particles in 1D

In this case, there are no constraints for the particles. The potential energy $u=0$ so that

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi}{dx^2} - E \Psi = 0.$$

This gives

$$\psi = A e^{-ikx} + B e^{ikx},$$

where $k = \sqrt{2mE} / \hbar = p / \hbar$ (note $E = p^2 / 2m$). The final solution is

$$\Psi_t(x,t) = A e^{-i(\omega t + kx)} + B e^{-i(\omega t - kx)},$$

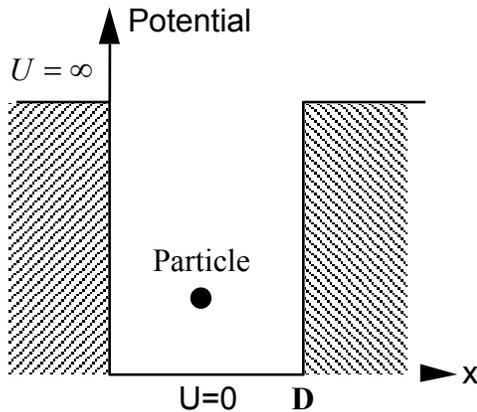
in which the first term corresponds to negative-direction propagation, the second term is positive-direction wave. Please also recall problem 2.5 in assignment 2.

2.3.2 Quantum well

Consider the general case of a particle in a one-dimensional potential well, which can be, for example, an electron subject to an electric potential field as shown in the figure. This is actually the model for thin films. The steady-state Schrödinger equation for the particle in such a potential profile is

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi}{dx^2} - E \Psi = 0 \quad (0 < x < D); \quad \Psi = 0 \quad (x < 0 \text{ or } x > D).$$

Note: for $u \rightarrow \infty$ ($x < 0$ or $x > D$), only $\Psi = 0$ can satisfy the Schrödinger equation.



Same as the free particles, the solution for first equation is still

$$\Psi = Ae^{-ikx} + Be^{ikx},$$

$$\text{where } k = \sqrt{\frac{2mE}{\hbar^2}} = \frac{\sqrt{2mE}}{\hbar} = \frac{p}{\hbar}.$$

The general boundary conditions are the continuity of the wave functions and their first derivatives at the boundaries. The latter derives from the continuity of particle flux at the boundary. For the current problem, the continuity of the first derivatives is not required because the wavefunction at the boundaries are already known to be zero. With the continuity of the wave function at $x=0$ and $x=D$, we have

$$\begin{cases} x=0 & A+B=0 \\ x=D & A \exp[-ikD] + B \exp[ikD] = 0 \end{cases}$$

Above equations yield

$$A(e^{-ikD} - e^{ikD}) = 0.$$

Noticing $A \neq 0$, we obtain $\sin(kD)=0$. Thus

$$k_n D = n\pi \quad (n=0, \pm 1, \pm 2 \dots)$$

or

$$D \sqrt{\frac{2mE_n}{\hbar^2}} = n\pi, \text{ energy eigen value } E_n = \frac{1}{2m} \left(\frac{\pi \hbar n}{D} \right)^2.$$

The material wave function inside the potential well is

$$\Psi_n = -2iA \sin\left(\frac{n\pi x}{D}\right),$$

which is identical with the previous results of standing wave.

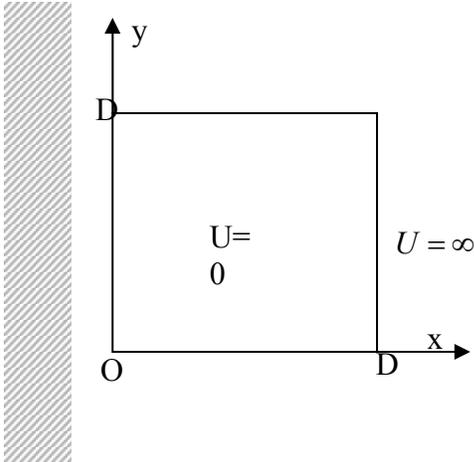
Note: for n , obviously it is not zero. Since solutions of negative n values are equivalent to those of positive values, here we can just let $n=1, 2 \dots$

Normalization is still used to calculate A , i.e.

$$\int_0^D \Psi_n^*(x) \Psi_n(x) dx = 1.$$

Finally we get $A = i\sqrt{\frac{1}{2D}}$.

2.3.2 Particle in a 2D box



We establish a coordinate system as shown in the schematic above. Clearly, outside the potential well, we have $\Psi = 0$ because $U = \infty$. We thus focus on the solution inside the potential well. The Schrödinger equation inside the well ($U=0$) is

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} \right) - E\Psi = 0.$$

We still use the separation-of-variables technique. Assuming $\Psi(x, y) = X(x)Y(y)$ and substituting into the above equation leads to

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{2mE}{\hbar^2} = 0.$$

In the above equation, the first term depends on x and the second term on y. The third term is a constant. This leads to the requirement that both the first and the second term must each be a constant. Since E is positive, we can prove neither of the first two terms can be positive. Errors will occur if we let one of them be positive. Thus, we write

$$\frac{1}{X} \frac{d^2 X}{dx^2} = -k_x^2, \quad \frac{1}{Y} \frac{d^2 Y}{dy^2} = -k_y^2.$$

The solution for X is

$$X(x) = A \sin(k_x x) + B \cos(k_x x).$$

To satisfy the boundary condition that $\Psi = 0$ at $x=0$ and $x=D$, we must have $X=0$ at $x=0$ and $x=D$. Applying these boundary conditions, we see that

$$k_x = \frac{n\pi}{D} \quad (n = 1, 2, 3, \dots).$$

Similarly, $k_y = \frac{l\pi}{D} \quad (l = 1, 2, 3, \dots).$

Thus

$$E_{\ell n} = \frac{(\ell^2 + n^2)\pi^2 \hbar^2}{2mD^2} \quad (\ell, n=1, 2, \dots)$$

and

$$\Psi_{\ell n} = C_{\ell n} \sin\left(\frac{n\pi x}{D}\right) \sin\left(\frac{\ell\pi y}{D}\right).$$

We can further determine the constant $C_{\ell n} = 2/D$.

For quantum numbers: n, l, \dots

$$(1) \quad n=1, \quad l=2, \quad E_{12}, \quad \Psi_{12}$$

$$(2) \quad n=2, \quad l=1, \quad E_{21}, \quad \Psi_{21}$$

Note: $E_{12} = E_{21}$, but Ψ_{12} and Ψ_{21} are flipped in the x and y directions from

$$\Psi_{\ell n} = C_{\ell n} \sin\left(\frac{n\pi x}{D}\right) \sin\left(\frac{\ell\pi y}{D}\right).$$

These states that have different wavefunctions but the

same energy are said to be degenerate. The degeneracy of an energy state is the number of wavefunctions having the same energy.

2.3.3 Electron spin & Pauli exclusion principle

Each wavefunction obtained in the previous sections represents a possible quantum mechanical state at which a particle can exist under the given potential. The solution of the Schrödinger equation, however, does not tell the whole story on the quantum state of a particle. For example it cannot distinguish the spin of particles. For electrons, corresponding to each wavefunction obtained from the Schrödinger equation, there are two quantum states (or two relativistic wavefunctions), which are usually denoted by an additional quantum number s that can have the following values:

$$s = \frac{1}{2} \quad \text{or} \quad -\frac{1}{2},$$

where $s=1/2$ is called spin up and $s=-1/2$ is called spin down.

The **Pauli exclusion principle** says that each quantum state can be occupied by at most one electron.

2.3.4 Other potentials

(1) Harmonic oscillator

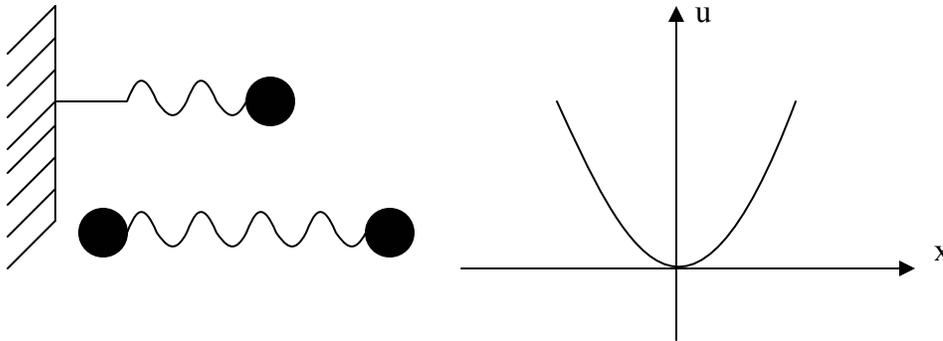
From $F = -Kx = -\frac{du}{dx}$, we have $u = \frac{1}{2}Kx^2$ (shown in the figure). The Schrödinger

equation becomes

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi}{dx^2} + \left(\frac{Kx^2}{2} - E\right)\Psi = 0,$$

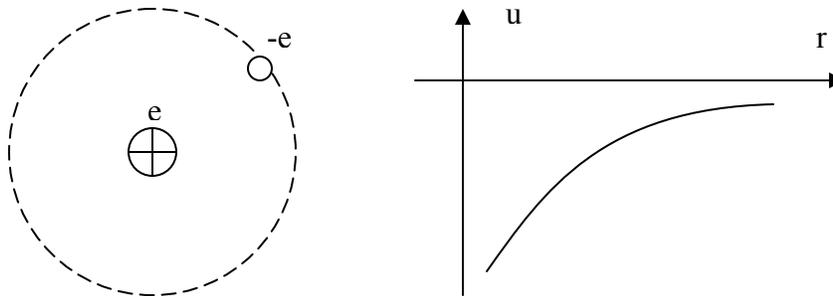
which gives $E_n = h\nu(n + 1/2); \nu = \frac{1}{2\pi} \sqrt{\frac{K}{m}}$.

Note: the zero point (intrinsic) energy $\frac{h\nu}{2}$ is required by the Heisenberg uncertainty principle ($\Delta t \Delta E \geq \frac{\hbar}{2}$). It shows that the oscillator has some energy even at the rest.



Note: to calculate the intrinsic frequency for diatomic molecules with two atoms of mass m_1 and m_2 , the reduced mass should be used,

$$m = \frac{m_1 m_2}{m_1 + m_2}.$$



Now consider the case that an electron moves around the nucleus, which is assumed to be stationary. The interaction between the nucleus and the orbiting electron is governed by the Coulomb force

$$F = -\frac{1}{4\pi\epsilon_0} \frac{e^2}{r^2} = -\frac{du}{dr},$$

where $\epsilon_0 = 1.124 \times 10^{-10} / 4\pi$ [$C^2 / (m^2 N)$] is the electrical permittivity of the vacuum. It yields

$$u(r) = -\frac{e^2}{4\pi\epsilon_0 r}.$$

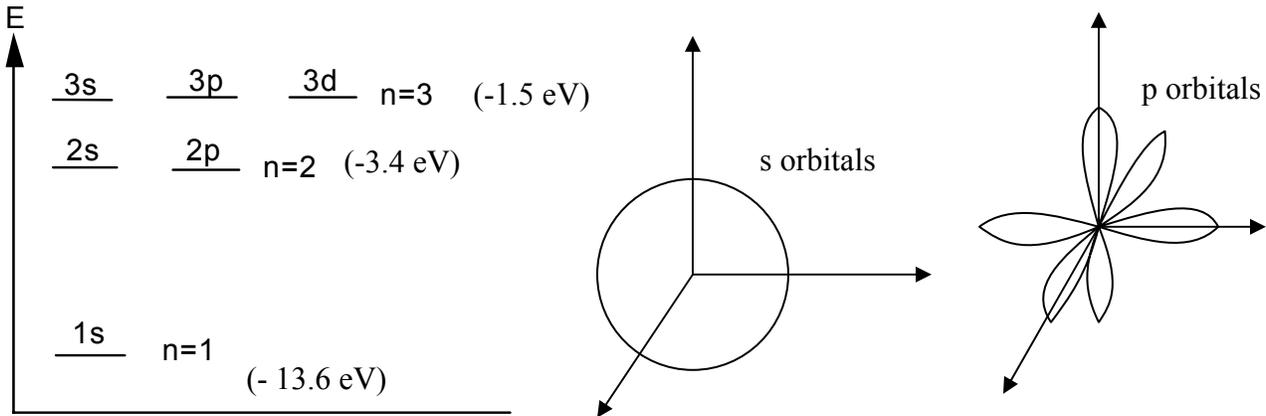
Using separation of variables, we assume $\Psi_{n\ell m} = R_{n\ell}(r)Y_\ell^m(\theta, \varphi)$. The allowable energy levels of the electron-nucleus system are

$$E_n^{el} = -\frac{Mc_1^2}{2\hbar^2 n^2} = -\frac{13.6 \text{ eV}}{n^2} \quad (n \geq 1, n \geq \ell + 1 \text{ and } |m| \leq \ell, \ell = 0, 1, 2, \dots)$$

in which M is the electron mass. For quantum number $n=1$, we have $n=1, \ell=0, m=0$, $\Psi_{100, s}$ (two quantum states determined by spin), 1s orbital.

$$n=2 \rightarrow \begin{cases} l=0 & m=0 & 2s \text{ orbital} \\ l=1 & m=-1,0,1 & 2p \text{ orbital} \end{cases}$$

Basically there are four wave functions for $n=2$ and totally eight quantum states. The energy levels of different orbitals are drawn in the following figure. The spherical s orbital and dumb-bell p orbitals are also presented here. The degeneracy is determined by $g=2n^2$.

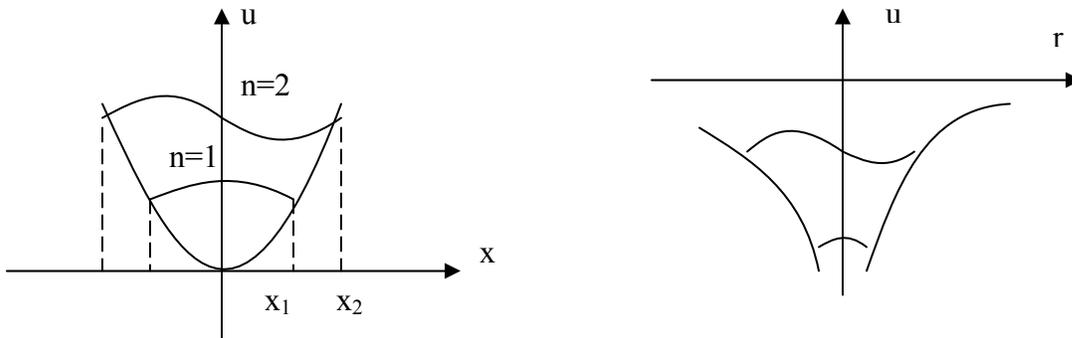


Note: the energy gap between different n values is much larger than the thermal energy ($k_B T \sim 26$ meV) and it is almost impossible to thermally excite electrons to a higher n level. A stable element is obtained only if all the orbitals for the highest n are completely filled, such as He.

Now one may wonder why the energy eigenvalue E has different relationships with n . Here we will give a simple argument, without solving the Schrodinger equation, to show that this is indeed the case. My argument is based on the requirement of forming standing waves in given potential. For harmonic oscillators, as shown in the following figure, the standing waves inside the potential, assuming at the boundaries, wavefunction is zero, give

$$x_n = \frac{\lambda n}{4}$$

$$\text{Kinetic energy is } KE = p^2 / 2m = \frac{1}{2m} \left(\frac{h}{\lambda}\right)^2 = \frac{1}{2m} \left(\frac{hn}{4x_n}\right)^2$$



Equating the kinetic and potential energies (sort of equipartition of energy, but only a hand waving argument), we obtain

$$\frac{K}{2} x_n^2 = \frac{1}{2m} \left(\frac{hn}{4x_n} \right)^2$$

or

$$x_n^2 = \left(\frac{1}{Km} \right)^{1/2} \frac{nh}{4}.$$

Thus

$$E \propto Kx_n^2 = \frac{nh}{4} \sqrt{\frac{K}{m}},$$

which is a linear functions of n, similar to what we obtain from solving Schrodinger equation. Similarly, we can deal with electrons moving around the nucleus but changing the potential energy to $u = \frac{e^2}{4\pi\epsilon_0 x_n}$. Finally we get $E \propto \frac{1}{n^2}$.

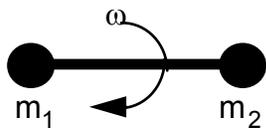
Rigid rotation

In classic mechanics, the angular momentum is expressed as $\vec{r} \times \vec{p}$. However, because of the uncertainty principle, it is difficult to give such an expression in quantum mechanics.

In classical mechanics, a quantity often used to describe the rotation is the moment of inertia. For a two-mass system rotating relative to its mass center, the moment of inertia is

$$I = \frac{m_1 m_2 r_0^2}{m_1 + m_2}$$

where r_0 is the effective separation between the two atoms.



The energy eigenvalues are

$$E_l = \frac{\hbar^2}{2I} \ell(\ell + 1) = hB\ell(\ell + 1) \quad (\text{for } |m| \leq \ell, \ell = 0, 1, 2, \dots).$$

Note: the discussion on harmonic oscillators and rigid rotors gives other useful information. For example, we can calculate the spring constant of an oscillator by

measured vibrational frequency, from $\nu = \frac{1}{2\pi} \sqrt{\frac{K}{m}}$.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 5

Quick review of Lecture 4

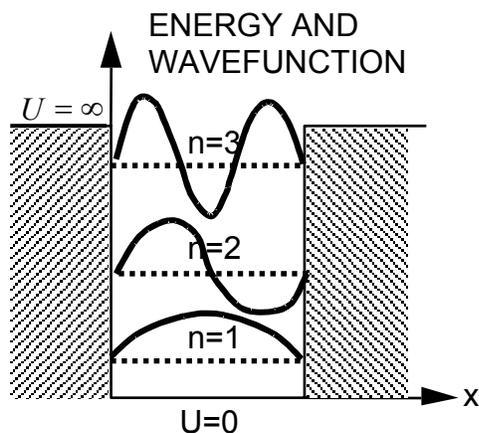
1. Free particles

The energy can be any values determined by the wavelength.

$$E = p^2 / 2m = \frac{(h / \lambda)^2}{2m} = \frac{\hbar^2 k^2}{2m}; k = 2\pi / \lambda, \hbar = h / 2\pi$$



2. Quantum well



Energy has discrete levels, and we have one quantum number n .

$$E = \frac{h^2}{8m} \frac{n^2}{D^2} \quad (n=1,2,\dots)$$

For 2D constraints, we have two quantum numbers n and l . In the discussions, the conception of “degeneracy” is introduced.

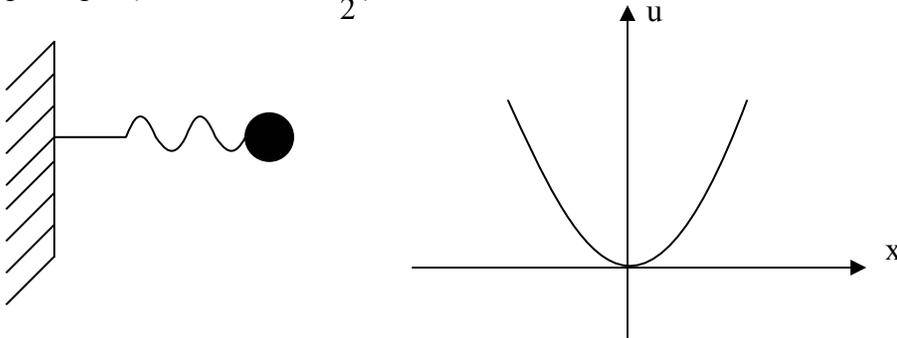
3. Spin

For electrons, we have talked about $s = \pm \frac{1}{2}$, where $s=1/2$ is called spin up and $s=-1/2$ is called spin down.

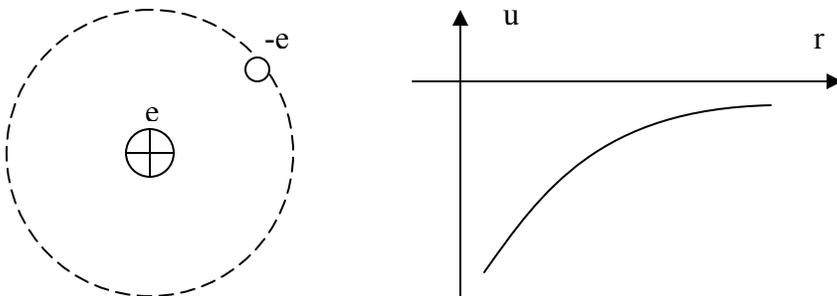
4. Harmonic oscillator

$$E_n = h\nu(n + 1/2); \nu = \frac{1}{2\pi} \sqrt{\frac{K}{m}}$$

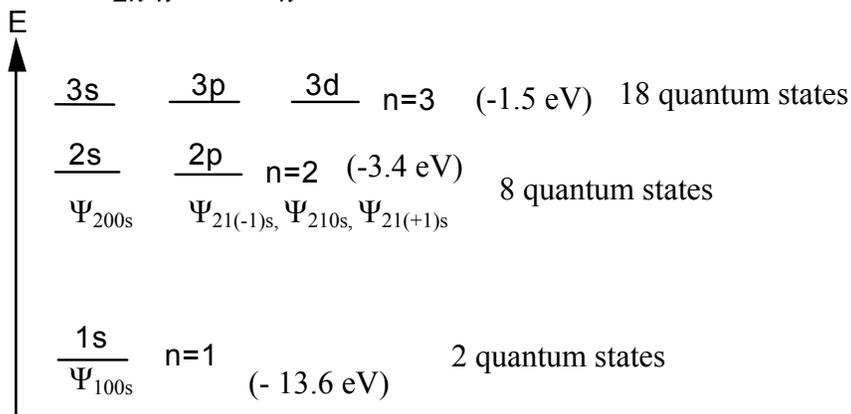
Note: the zero point (intrinsic) energy $\frac{h\nu}{2}$ is required by the Heisenberg uncertainty principle ($\langle \Delta t \rangle \langle \Delta E \rangle \geq \frac{\hbar}{2}$).



5. Electrons moving around the nucleus



$$E_n^{el} = -\frac{Mc_1^2}{2\hbar^2 n^2} = -\frac{13.6 \text{ eV}}{n^2} \quad (n \geq 1, n \geq \ell + 1 \text{ and } |m| \leq \ell, \ell = 0, 1, 2, \dots)$$



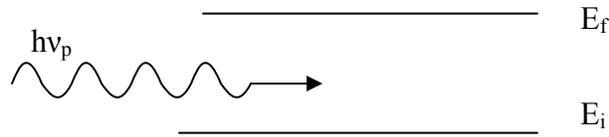
The corresponding Ψ_{nlms} are marked in the figure for some energy levels. The degeneracy follows $g(n)=2n^2$.

Note: As electrons number goes up, the orbit will split. The energy of 3d is lifted up above 4s because of the electron-electron interaction. For the element potassium (K), it has 19 electrons but the n=3 energy levels are not totally filled and one electron goes to the 4s orbit.

2.3.5 Energy Quantization observation

Absorption or emission of photon happening only

$$E(\text{photon}) = h\nu_p = E_f - E_i$$



The allowable energy levels of the electron-nucleus system (hydrogen atom) are

$$E_n^{el} = -\frac{13.6 \text{ eV}}{n^2}.$$

The emission occurring between $n=1,2$ is

$$h\nu_p = -13.6 \text{ eV} \left(\frac{1}{2^2} - \frac{1}{1^2} \right) \sim 10 \text{ eV}.$$

The corresponding wavelength is

$$\lambda_p = c / \nu_p.$$

Sometimes, we also use the wave number as

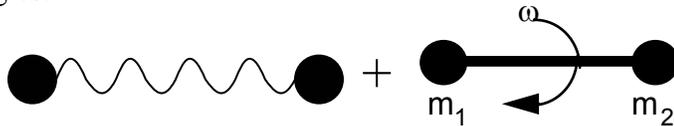
$$\frac{1}{\lambda_p} = \frac{\nu_p}{c} [\text{cm}^{-1}].$$

These units are used interchangeably and you should be able to do the conversion yourself. One good number to remember is that 1eV is 1.24 μm .

Now we are in a position to discuss the total energy of an atom or molecule. The total energy can be approximated as the summation of translational, vibrational, rotational, and electronic energies:

$$E^{tot} = E^{trans} + E^{el} + E^{vib} + E^{rot}.$$

We talk about the translational energy as a particle in a box. To simplify, for hydrogen molecules we neglect other effects and only consider the vibrational and rotational energies.



If the emission and absorption occur between two energy levels of vibration, we have

$$E_n = h\nu(n + 1/2).$$

Thus

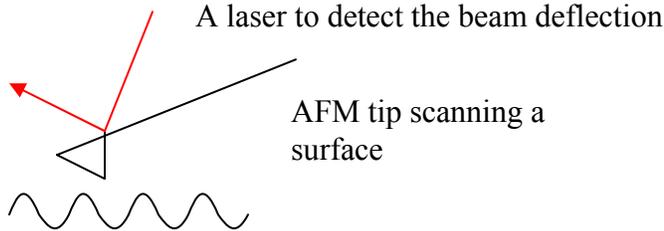
$$h\nu_p = h\nu(n_f - n_i); \nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}},$$

where the reduced mass is $m = \frac{m_1 m_2}{m_1 + m_2} = \frac{m_1}{2}$ ($m_1 = m_2$), and the selection rule states

$n_f - n_i = \pm 1$. Positive for absorption and negative for emission. We can measure the vibrational frequency from which, to deduce the spring constant. For hydrogen, the

vibrational wavelength is 2.3 micron, corresponding to a $k \sim 500$ N/m. This is a large spring constant.

Note: The large spring constant between atoms is one reason we can use atomic force microscopy to measure the topology of a surface without damaging the surface. The cantilever of the AFM has a spring constant much smaller than that of the spring constant between atoms. When a sharp tip built on the cantilever scans over a solid surface, because the spring constant of the beam is much lower than that of the atomic vibration, the beam will deform instead of deforming the atoms on the surface. The beam deformation can be further measured by a laser and the topography-scanning resolution can reach nm level and even atomic resolution.



For rotational energy, we have 2 degrees of freedom. The energy eigenvalues are

$$E_l = \frac{\hbar^2}{2I} l(l+1) = hB l(l+1) \quad (\text{for } |m| \leq l, l=0,1,2, \dots).$$

$$h\nu_p = hB[l_f(l_f+1) - l_i(l_i+1)],$$

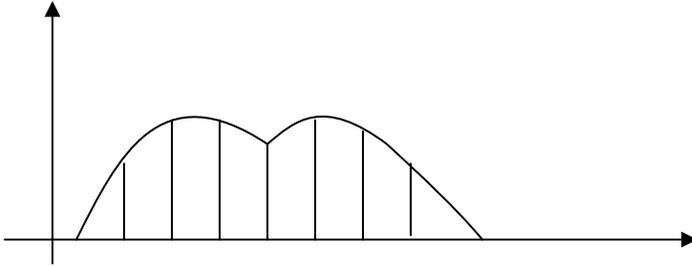
where $B=1.8e12$ Hz. Similarly we have $l_f - l_i = \pm 1, 0$. As an example, $l_f - l_i = 1$ (absorption) gives $\nu_p = 2B(l_i + 1)$. The normal wavelength for rotational energy is around $100 \mu\text{m}$ (far infrared regime), which is much larger than the emission wavelength at the room temperature ($10 \mu\text{m}$).

The energy bands are shown in the following figure, in which rotational energy accounts for small energy sub bands.

$$\nu_p = \nu_{\text{vibration}} \pm 2B(l+1).$$



The degeneracy $g(l)=2l+1$. In the following figure, the density of bands increases with increasing l .

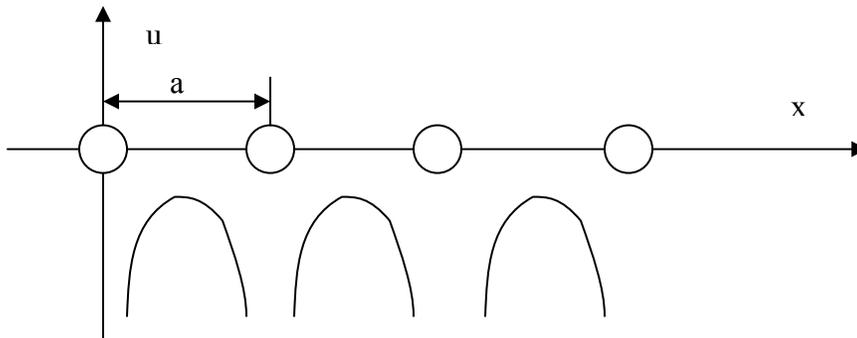


For CO_2 ($O=C=O$), three basic modes exist: symmetric stretching, asymmetric stretching, and bending. The absorption wavelength is around $10\ \mu\text{m}$, which is just the earth emission wavelength at the room temperature.

Note: the green housing effects occur when solar emission ($\sim 0.5\ \mu\text{m}$) can pass CO_2 and heat up the earth. Then the emission of the earth will be absorbed by CO_2 and the heat is trapped on earth, leading to global warming.

Chapter 3 Energy states in a solid

3.1.1 A crystal is periodic arrangement of atoms



Consider a one-dimensional lattice. The potential energies of adjacent atoms overlap and form a periodic potential field, i.e. $u(x+na) = u(x)$. Instead of the boundary conditions used for particles with constraints, now we will employ the periodicity to calculate the energy levels of electrons.

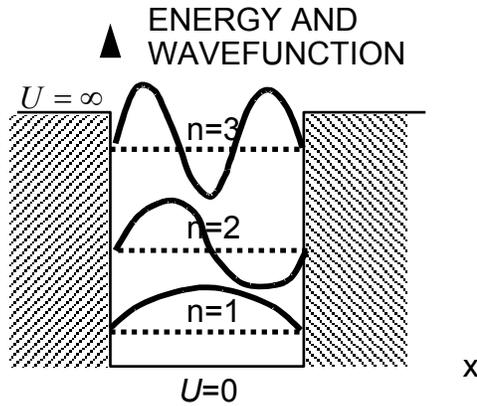
3.1.2 Possible approaches

First we will review the problems solved in Lecture 4.

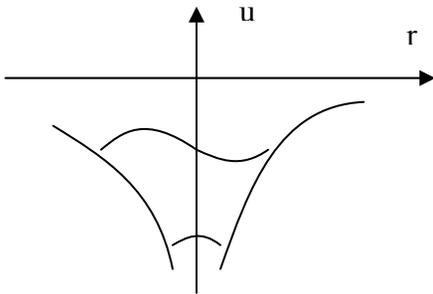
(1) Free electron $E = p^2 / 2m$



(2) Quantum well $E \propto n^2$

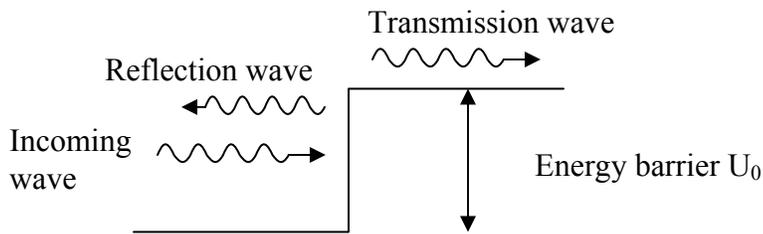


(3) Electron-nucleus system (hydrogen atom) $E_n^{el} = -\frac{13.6 eV}{n^2}$

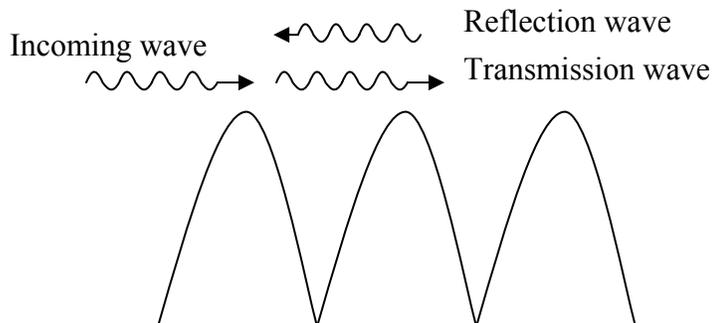


(4) Homework 2.5. At a potential step, there are reflections and transmissions of waves.

- { $E > U_0$, transmission wave Ψ_t propagates.
- { $E < U_0$, transmission wave Ψ_t decays from the interface.

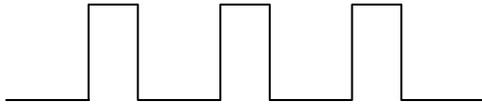


Later we will talk more about the interface influence. In the following figure, there will be interference effects of the reflection waves from different interfaces.

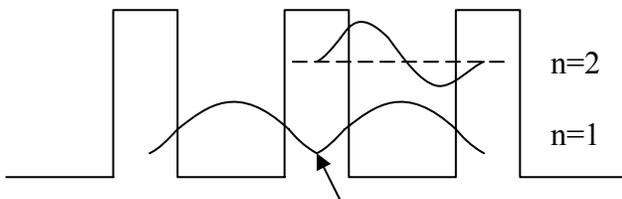
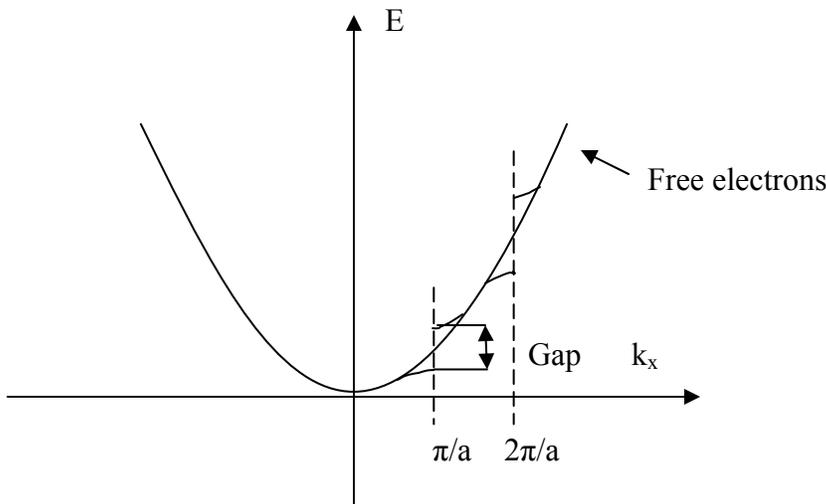


Now we will start solving the Schrödinger equation with periodic u :

(1) First we will consider periodic quantum wells with finite depth, and electrons with energy larger than the barrier function as free electrons. Basically we approximate the potential field as rectangular wells.



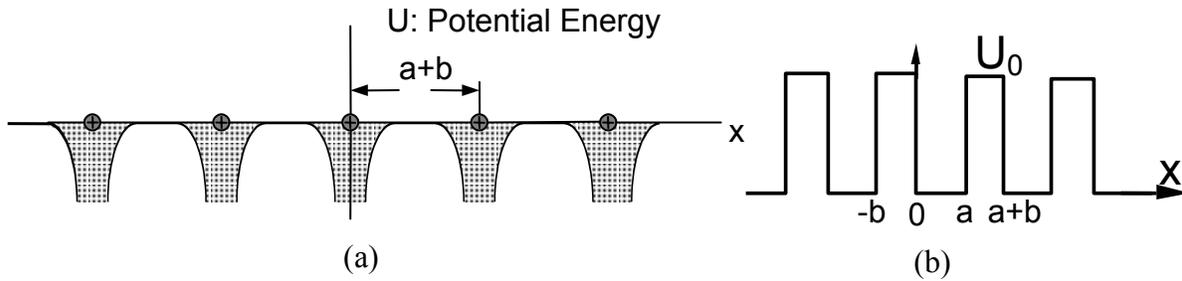
Compared with a free electron, there will be small energy gaps on the energy curve for $k_x = n\pi/a$ ($n = \pm 1, \pm 2, \dots$). No electrons are allowed on the energy levels within the gaps. This is caused by the interference effect, which we will learn in details in chapter 5.



NOT allowed by Pauli exclusion principle

Now let us go back to the quantum wells. Due to the finite depth, the wave functions will not be zero at the boundary like standing waves. Instead, they will decay exponentially from the interface. However, this indicates two wavefunctions overlap in the middle of the barrier, which conflicts with the Pauli exclusion principle. The waves will split in this situation. We will talk more about this later.

(2) Kronig-Penney model



Here we will solve the Schrödinger equation for one period

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi + (U - E)\Psi = 0 \quad (0 < x < a+b).$$

Recall homework 2.5. We have one positive-direction wave and one negative-direction wave in the same region. Two coefficients A and B need to be determined in

$$\Psi = Ae^{-iKx} + Be^{iKx}.$$

For $[0, a]$ and $[a, a+b]$ regions, we have totally four unknown coefficients. The continuity boundary condition at $x=a$ only gives two equations

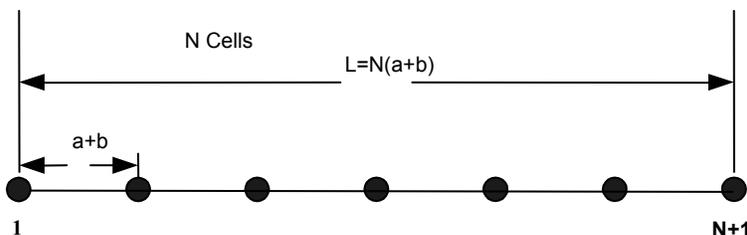
$$\Psi_1(x=a) = \Psi_2(x=a), \quad \Psi_1'(x=a) = \Psi_2'(x=a).$$

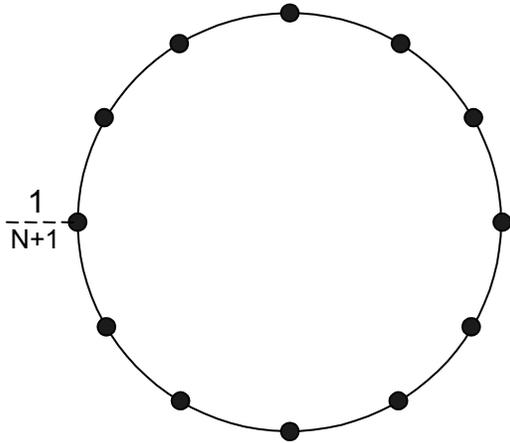
The other two equations come from the potential periodicity. Because the potential is periodic, the wavefunctions between different periods are related through the Bloch theorem, i.e.

$$\Psi[x + (a+b)n] = \Psi(x)e^{ikn(a+b)},$$

Therefore, we finally have four equations to determine the constants. Please note this k is different from K in the Ψ expression.

We now determine the value of the wvector k in the Bloch theorem, using the **Born-von Karman** periodic boundary condition. The Born-von Karman boundary condition deals with the end points of a crystal. Ordinarily, we would think that the two end points are different from the internal points. For many applications, however, distinguishing the boundary points from the internal points is not necessary, because a crystal usually has a tremendous number of lattice points (this is not true for quantum wells, quantum wires, and quantum dots). The Born-von Karman boundary condition requires that the wave functions at the two end points be equal to each other, i.e., the two end points are overlapped to form a lattice loop as shown in the figure.





First we have

$$\Psi[x + N(a + b)] = \Psi(x).$$

Using Bloch's theorem, it can be written as

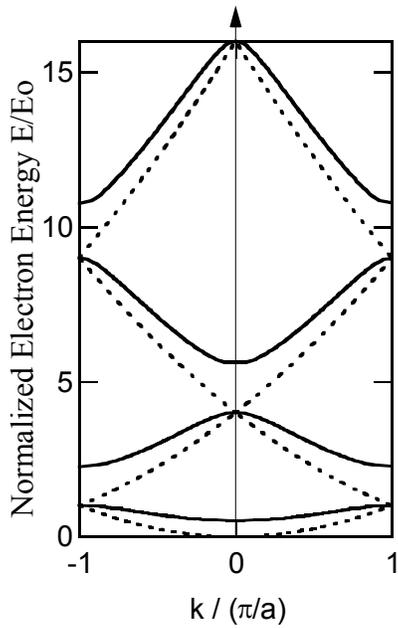
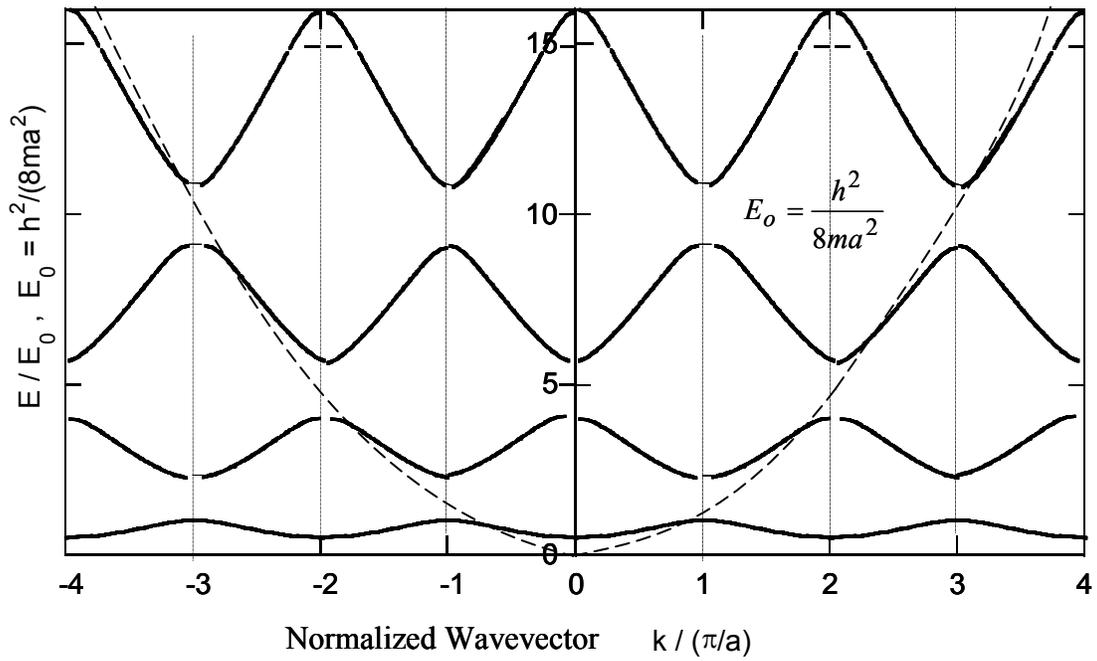
$$\Psi(x) = \Psi(x) \exp[ikN(a + b)],$$

which yields

$$k = \frac{2\pi n}{N(a + b)} = \frac{2\pi n}{L} \quad (n=0, \pm 1, \pm 2, \dots),$$

where L is the length of the crystal.

To get a better idea of the solution form, now we simply let b approach zero. The following figure shows that for each wavevector k , there are many possible values for the electron energy E . These values form quasi-continuous bands as a function of k (because k itself is quasi-continuous as to be discussed latter). Because both the wavefunctions and the eigen energy for the states correspond to the wavevectors k and $[k+m(2\pi/a)]$ are identical, these are actually the same quantum state and should be counted only once. Thus, rather than plotting the energy eigenvalues for all the wavevectors, we can plot them in one period, as shown in the subsequent figure. This way of representation is called the **reduced-zone** representation. Often, only half of the band, $[0, \pi/a]$, needs to be drawn because the band is symmetric for both positive and negative wavevector values. The relationship between the energy and the wave vector is the dispersion relation.



2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 6

Quick review of Lecture 5

In the last lecture, we approximate the potential field as rectangular wells in the crystal. From periodicity, the Bloch theorem gives additional equations

$$\Psi[x + (a + b)n] = \Psi(x)e^{ikn(a+b)},$$

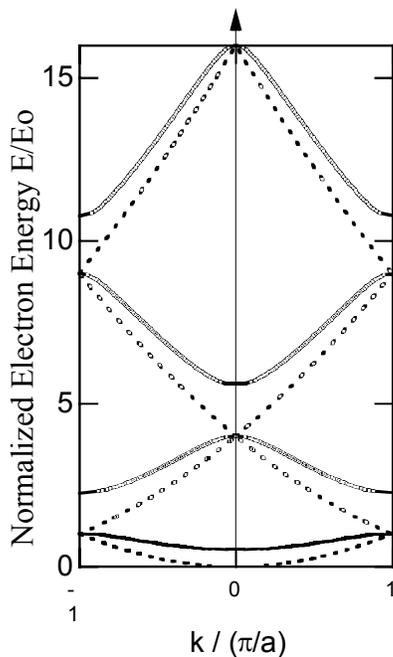
which is used to determine the coefficients in $\Psi = Ae^{-iKx} + Be^{iKx}$. We have also determined the value of the wvector k in the Bloch theorem, using the **Born-von Karman** periodic boundary condition $\Psi[x + N(a + b)] = \Psi(x)$. This yields allowed k values as

$$k = \frac{2\pi n}{N(a + b)} = \frac{2\pi n}{L} \quad (n=0, \pm 1, \pm 2, \dots),$$

where L is the length of the crystal.

For each k_n , there are two quantum states denoted by $\Psi_{k_n, s}$ (spin up, spin down). When n goes from 1 to N , k_n varies from 0 to $2\pi/a$. Therefore, in the following E-k figure we totally have $2N$ quantum states $\Psi_{k_n, s}$.

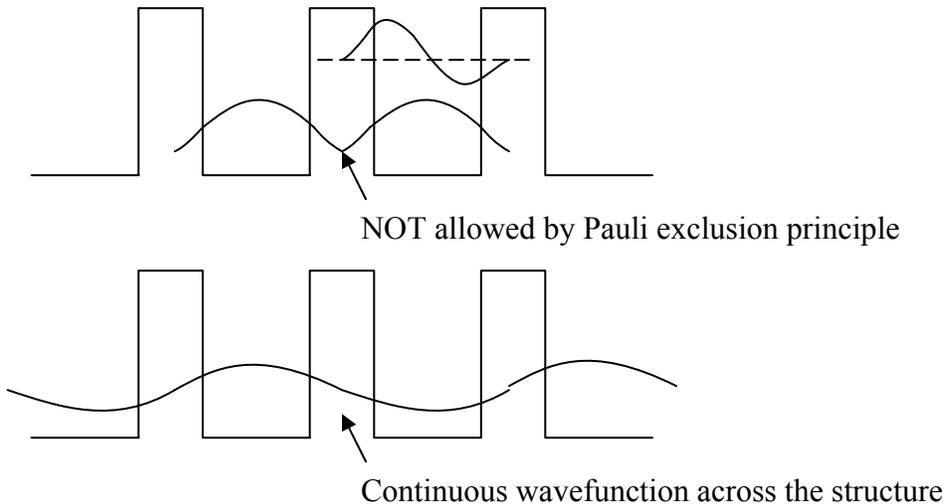
Note: (1) For big crystals, N is very big (on the magnitude of 10^{23}) and $\Delta k = \frac{2\pi}{L}$ is also small. The following curve can be regarded as quasi-continuous. (2) The **Born-von Karman** periodic boundary condition is no longer valid when N becomes very small in nanomaterials.



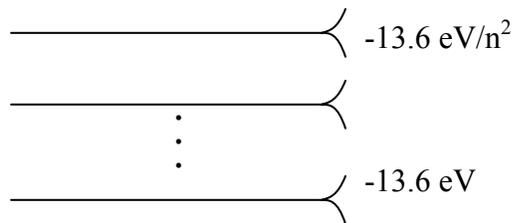
3.1.3 Consequences of solid energy levels we just obtained:

(1) Electrons wave function extends through the whole crystal, they belong to all atoms collectively.

Recall the splitting of waves discussed in last lecture. According to Pauli's exclusion principle, the wavefunctions of adjacent wells cannot overlap. A continuous wave extending through the whole structure will be formed in this situation. The wavefunction no longer corresponds to an individual atom.

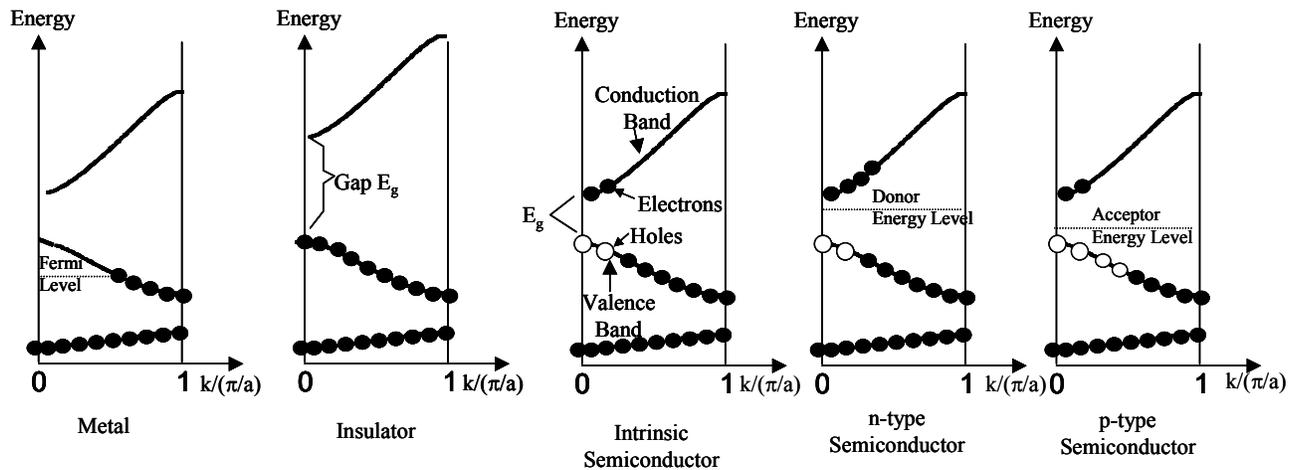


Similar argument also exists for atomic energy levels. When two atoms become closer, the overlap of electron wave functions will cause band split.



(2) Filling of electrons

As mentioned before, in E-k figure every band (k changes for $2\pi/a$) can accommodate $2N$ quantum states. At zero temperature, the filling rule for the electrons is that they always fill the lowest energy level first, as required by thermodynamics. If one atom only has one electron, the band is half filled since there are only N valence electrons in this case, as shown in the next figure. The topmost energy level that is filled with electrons at zero Kelvin is called the **Fermi level**. The electron energy and momentum can be changed (almost) continuously within the same band because the separation between successive energy levels is small. Thus, these electrons can flow freely, making the materials good electrical conductors, which is the case for metals.

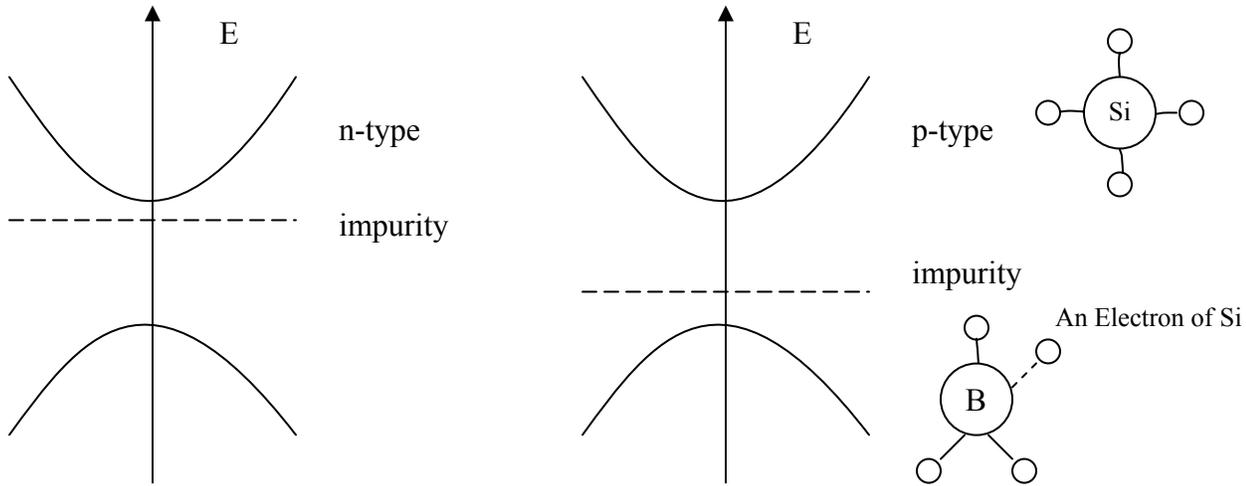


If the valence electrons exactly fill one or more bands, leaving others empty, the crystal will be an insulator at zero temperature and can be an insulator or a semiconductor at other temperatures depending on the value of the energy gap. If a filled band is separated by a large energy gap (>3 eV) from the next higher band, one cannot change the energy and the momentum of an electron in the filled band easily, that is, these electrons cannot move freely and the materials are insulators. A semiconductor is essentially similar to an insulator. The difference between them is that the gap between the filled and the empty bands for a semiconductor are not so large (<3 eV), so that some electrons have enough thermal energy (at room temperature $k_B T = 0.026$ eV) to jump across the gap to the empty band above (called conduction band), and these electrons can conduct electricity (these are called intrinsic semiconductors). The unoccupied states left behind also leave room for the electrons in the original band (called valence band) to move. It turns out that the description of the motion of these electrons is equivalent to thinking that the vacant states move as positive electrons, or holes. The energy of these holes is a minimum at the peak in the valence band and increases as the electron energy becomes more negative.

Note: (1) Here we only talk about 1D case, for 3D structure we have different band shapes in different direction and the wavefunction changes to $\Psi(k_x, k_y, k_z)$. Later we will talk about the electron filling in a 3D structure. (2) The shape of bands can be affected by heavily doping. (3) The lowest band starts from a nonzero energy, which is the consequence of the uncertainty principle.

Impurities are added to most semiconductors and these impurities have energy levels somewhere within the band gap, some are close to the bottom edge of the conduction band or top edge of the valence band (or band edge). The electrons in the impurity levels can be thermally excited to the conduction band if their level is close to the bottom of the conduction band, creating more electrons in the semiconductor than holes. Such semiconductors are called n-type and the impurities are called donors. A typical example is P. Similarly, if the impurity energy level is close the valence band edge, electrons in

the valence band can be excited to the impurity levels, leaving more empty states or holes behind. Such semiconductors are called p-type and the impurities are called acceptors. For instance, a B atom has three outer core electrons; it will catch one more electron from Si atoms.



In semiconductors, the moving charge carriers normally are near the minima or maxima of a band, where $\frac{\partial E}{\partial k} = 0$. Taylor's expansion gives

$$\begin{aligned}
 E(k) &= E(k_m) + \frac{\partial E}{\partial k} \Big|_{k_m} (k - k_m) + \frac{1}{2} \frac{\partial^2 E}{\partial k^2} \Big|_{k_m} (k - k_m)^2 \\
 &= E(k_m) + \frac{1}{2} \frac{\partial^2 E}{\partial k^2} \Big|_{k_m} (k - k_m)^2 \\
 &= E(k_m) + \frac{1}{2} \frac{\hbar^2}{m^*} (k - k_m)^2,
 \end{aligned}$$

where the effective mass is defined as $m^* = \frac{\hbar^2}{(\partial^2 E / \partial k^2) \Big|_{k_m}}$. In differential geometry, the

term $1 / \partial^2 E / \partial k^2$ is just the curvature. Thus, effective mass is proportional to local curvature at band maxima or minima. For electrons close to the minima of the conduction band, we have

$$E - E_c = \frac{\hbar^2 k^2}{2m^*},$$

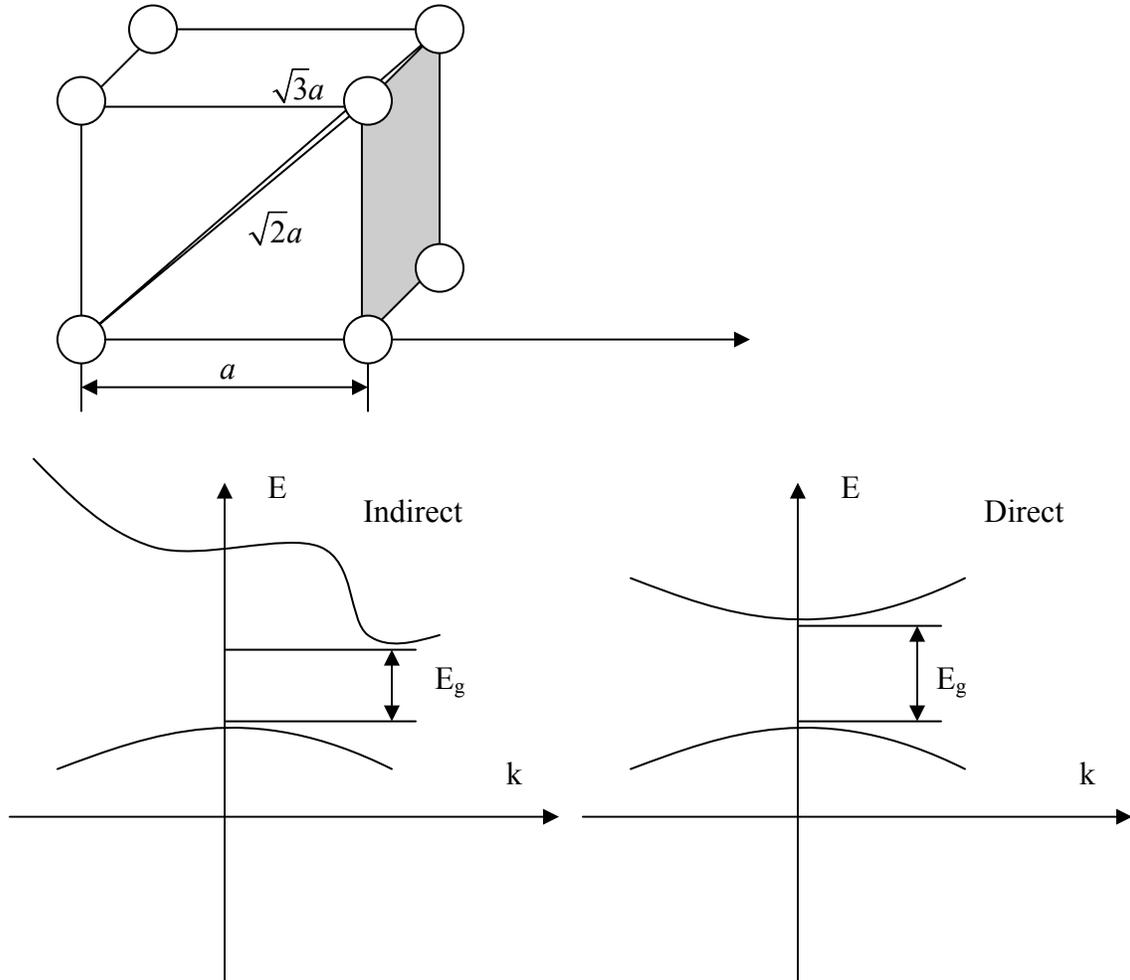
from which we can clearly see the meaning of effective mass by comparing with the energy of free electrons

$$E = \hbar^2 k^2 / 2m.$$

In the above equation, $\hbar k$ represents the momentum of the free electron. The momentum of an electron in the crystal, however, should be calculated from the wavefunction using the momentum operator $-i\hbar\nabla\Psi$. Such a calculation would show that $\hbar k$ is not the momentum of the electron. Nevertheless, in many ways $\hbar k$ for a periodic potential behaves as the momentum of a free electron and thus it is called the **crystal momentum**.

3.2 Different directions of a crystal, different E(k) relation

In the real crystal, the band shapes normally differ a lot in different directions. In the following figure, the periods are a , $\sqrt{2}a$, and $\sqrt{3}a$ respectively in three directions.



In above figures, we demonstrate the idea of direct and indirect band gaps. In a direct band structure, both the minimum in the conduction band and the maximum in the valence band occur at the same location of k ($=0$ for the example given). A good example for direct band gap is GaAs with $E_g = 1.42$ eV, while Si is indirect gap semiconductor. Direct gap semiconductors are used in lasers, while Si is for microelectronics.

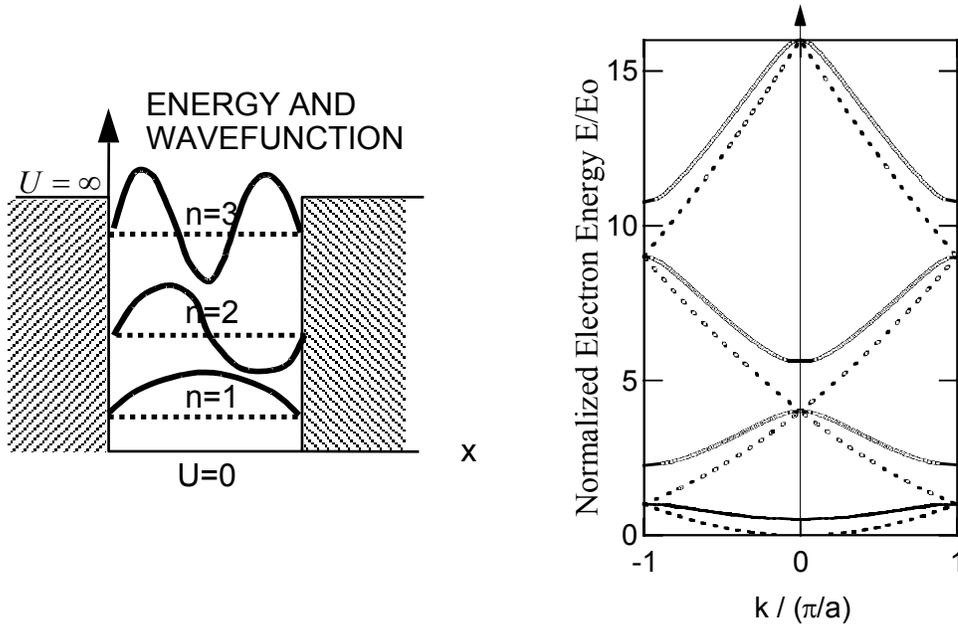
3.3 Lattice vibration and phonons

We have considered a harmonic oscillator for the H_2 molecule. The energy is expressed as

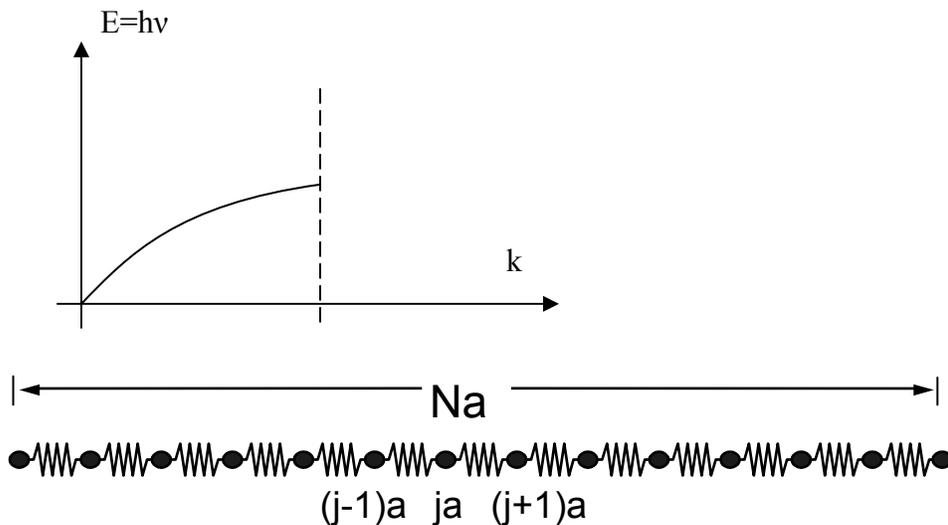
$$E_n = h\nu\left(n + \frac{1}{2}\right); \nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \quad (n=0,1,2,\dots)$$



Now let us think about N atoms in a one-dimensional chain. First recall the following cases. In the first figure, the energy is quantized, while band gap appears in the second figure.



Similarly, we can anticipate the single energy level of a diatomic molecule will split into a band in a lattice chain of atoms.



We make the following assumptions for the analysis. First, we consider only the interaction force between the nearest neighbors. Second, the interaction force between atoms is a harmonic force (which obeys Hook's law). This can be justified as we have done for harmonic oscillators. Now consider a typical atom j . The displacement of atom j from its equilibrium position x_j^0 is

$$u_j = x_j - x_j^o$$

The force acting on atom j comes from two parts. One is due to the relative displacement between atom $(j-1)$ and atom j , and the other is due to the relative displacement between atom j and $(j+1)$. The net force is then

$$F_j = K(u_{j+1} - u_j) - K(u_j - u_{j-1}).$$

Applying Newton's second law to atom j , we obtain

$$m \frac{d^2 u_j}{dt^2} = K(u_{j+1} - u_j) - K(u_j - u_{j-1}).$$

The above equation is a special form of the differential wave equation

$$m \frac{\partial^2 u}{\partial t^2} = Ka^2 \frac{\partial^2 u}{\partial x^2}.$$

which has a solution of the form $u \propto e^{-i(\omega t - kx)}$. Since a is not very small, we cannot use this solution directly. But this leads us to guess a wave type of solution as

$$u_j = A \exp[-i(\omega t - kja)].$$

Substituting the guessed solution into the discrete equation, we get

$$-m\omega^2 = K[e^{ika} + e^{-ika} - 2]$$

or

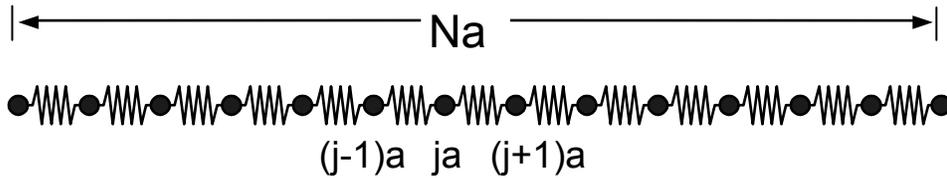
$$\omega = 2\sqrt{\frac{K}{m}} \left| \sin \frac{ka}{2} \right|.$$

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 7

In the last lecture, we have talked about atoms in a one-dimensional chain. We find the solution as

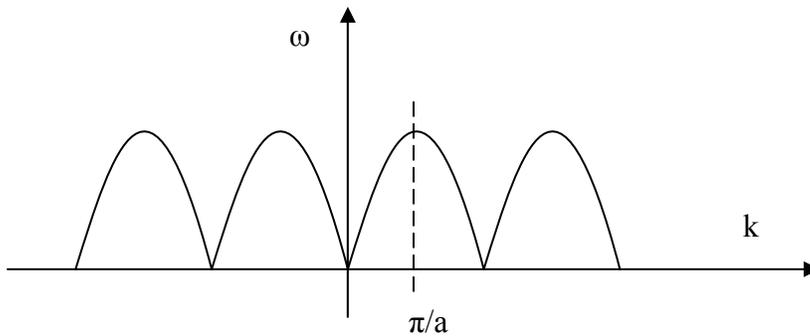
$$u_j = A \exp[-i(\omega t - kja)],$$

where the frequency is $\omega = 2\sqrt{\frac{K}{m}} \left| \sin \frac{ka}{2} \right|$.



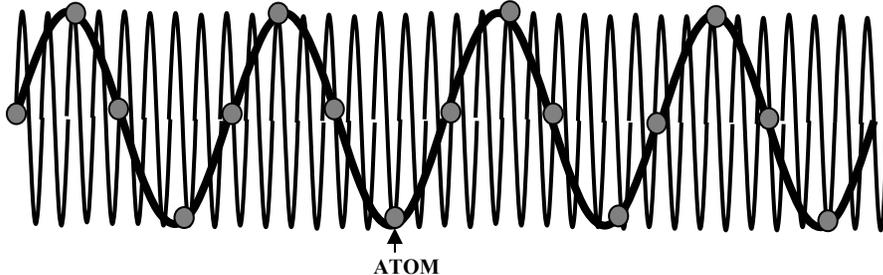
Note: When k approaches zero for large wavelength, the frequency becomes a linear function of the wavevector, i.e. $\omega \approx 2\sqrt{\frac{K}{m}} \frac{ka}{2} = ka\sqrt{\frac{K}{m}}$. We can calculate the sound

velocity by $v_{\text{sound}} = \frac{d\omega}{dk}$.



In last two lectures, we have derived the allowed k values as

$$k = \frac{2\pi n}{Na} \quad (n=0, \pm 1, \pm 2, \dots).$$



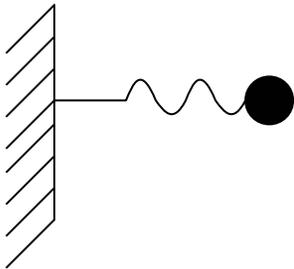
In the above figure, $k = \pi/a$ corresponds to $\lambda = 2a$. However, $k > \pi/a$ is meaningless in physics because there is no atoms vibrating between one period. Thus the allowable

wavevector for a lattice vibration is naturally confined to the first Brillouin zone ($|k| < \pi/a$). Therefore, we totally have N allowable wavevectors (also wavelength) in between $-\pi/a < k < \pi/a$. Each of these wavevectors corresponds to one mode of the vibration of the lattice. This mode is called a **normal mode**.

For a harmonic oscillator, quantum mechanics gives

$$E_n = h\nu(n + 1/2); \nu = \frac{1}{2\pi} \sqrt{\frac{K}{m}},$$

in which the frequency is consistent with classical mechanics.



Similarly, quantum mechanics gives the energy levels for a chain with N atoms as

$$E_n = h\nu(n + \frac{1}{2})$$

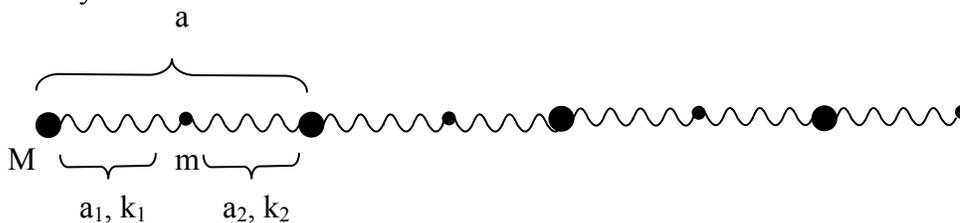
or $E_n = \hbar\omega(n + \frac{1}{2})$ ($n=0,1,2,3,\dots$),

in which $\omega = 2\sqrt{\frac{K}{m}} \left| \sin \frac{ka}{2} \right|$.

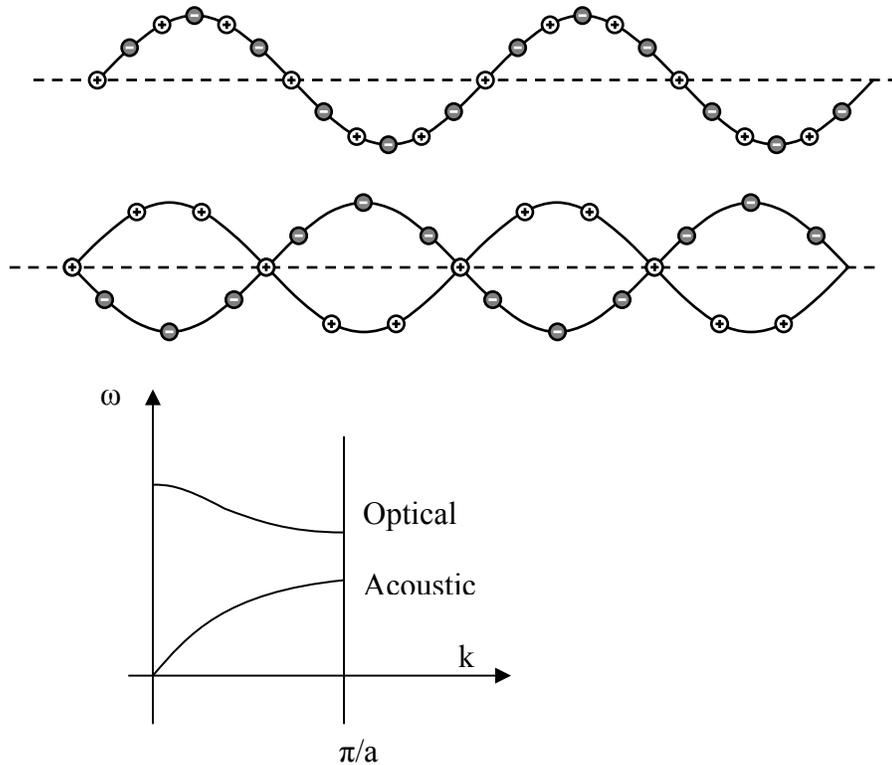
The basic vibrational energy quanta, $h\nu$, is called a phonon. Comparison between electrons, phonons, and photons:

- (1) Electrons obey the Pauli exclusion principle, which says that each quantum state can only have at most one electron. Photons and phonons are not limited by the Pauli exclusion principle. Each quantum state, which corresponds to one set of wavevectors, can have many phonons and photons.
- (2) Unlike electrons, phonons and photons at rest do not have mass though they have momentum and energy. They are also called fictitious particles since they are the quantization of the normal mode of a field.

3.3.3 Polyatomic lattice chain

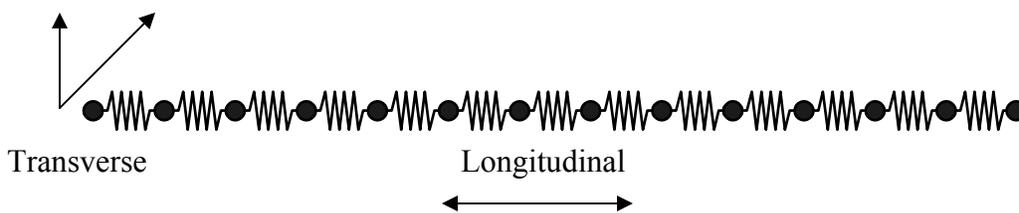


Now let us consider a chain with two atoms per period, in which two types of motions exist. In the first case, the adjacent two atoms vibrate in phase, while in the second figure the two atoms are moving out of phase. Clearly, the out-of-phase modes require more energy. Lower frequency (energy) branch is called the acoustic branch and the higher frequency one is called the optical branch, because the high frequency phonons in the optical branch can interact with electromagnetic waves more easily. In general, if there are m atoms in a basis and N lattice points in the chain, there are one acoustic branch with N acoustic modes, and $(m-1)$ optical branches with $(m-1)N$ optic modes.

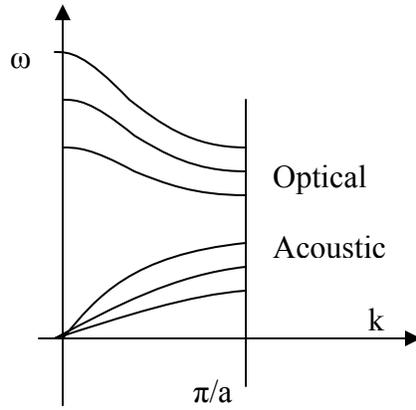


3.3.4 Phonons in 3D crystals

- a. Each direction is different
- b. Two transverse waves, one longitudinal wave



- c. For m atoms per basis, we have 3 acoustic waves and $3(m-1)$ optical waves.
- d. Each k_x, k_y, k_z represents a normal mode.
- e. The energy dispersion (E-k relationship) can be totally different in different directions, such as $a, \sqrt{2}a$, and $\sqrt{3}a$.



Note: (1) For the out-of-phase movement, both longitudinal and transverse modes exist. (2) The transverse waves normally have lower frequencies because less energy is required for the wave propagation. (3) Normally sound (pressure waves) are longitudinal waves; electromagnetic (EM) waves are transverse waves; acoustic waves can be either type. (4) In some cases two transverse waves can overlap in the figure. (5) At $k = \pi/a$, all curves should be horizontal.

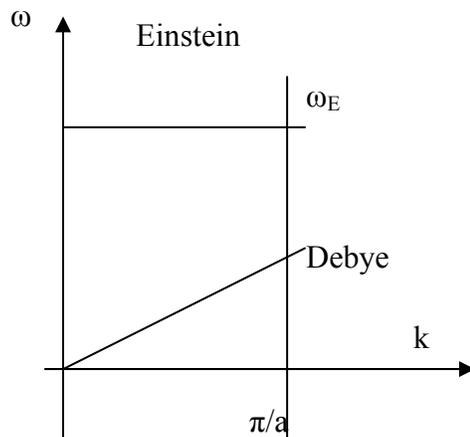
f. Approximation

(1) Debye approximation

Although the E-k curves for acoustic phonons are nonlinear, very often, the Debye approximation is used, which assumes a **linear** and **isotropic** relation between the frequency and the wavevector. This approximation is valid at low frequencies but is not a good approximation at high frequencies. In the very low frequency region, the lattice vibration carries the sound wave.

(2) Einstein approximation

The optical phonons are simplified as constant frequency ω_E .



3.1 Crystal structures

(1) Basic conceptions

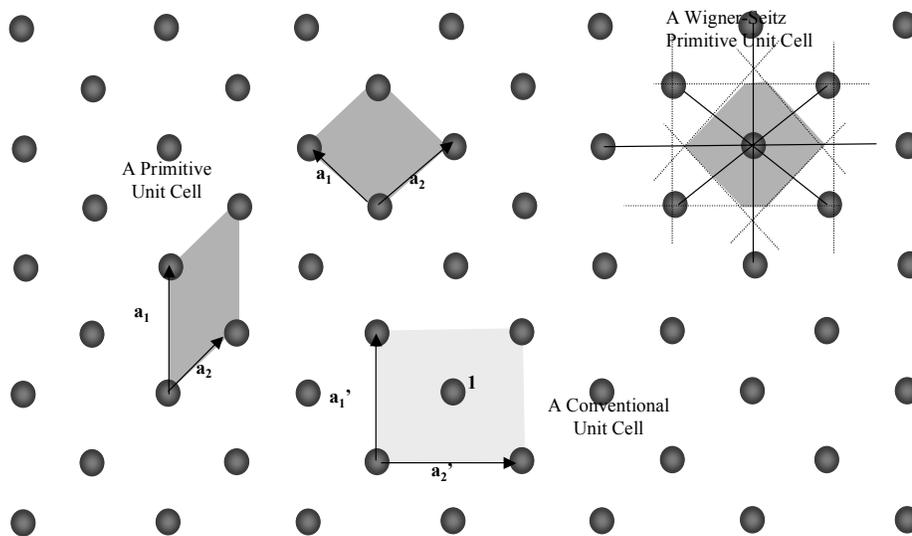
To form an actual crystal, a basis consisting of one or several atoms (or a molecule) is attached to each lattice point, i.e.

crystal = lattice + basis.

From a mathematical point of view, the location of each point can be described by a vector. Due to the periodic arrangement of lattice points, we can choose a basic set of vectors called the **primitive lattice vectors** to construct all other vectors in the lattice. In a three-dimensional lattice, \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{a}_3 are primitive lattice vectors if from any point, we could reach all other lattice points by a proper choice of integers through the following translation

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (n_1, n_2, n_3 \text{ cover all integers}).$$

The magnitudes of \mathbf{a}_1 , \mathbf{a}_2 , and \mathbf{a}_3 are called the lattice constants. The other set of vectors \mathbf{a}_1' and \mathbf{a}_2' are not primitive lattice vectors because we cannot use them to construct all other lattice points by a two-dimensional version of the above equation. For example we cannot reach point 1 through any linear integer combination of \mathbf{a}_1' and \mathbf{a}_2' .

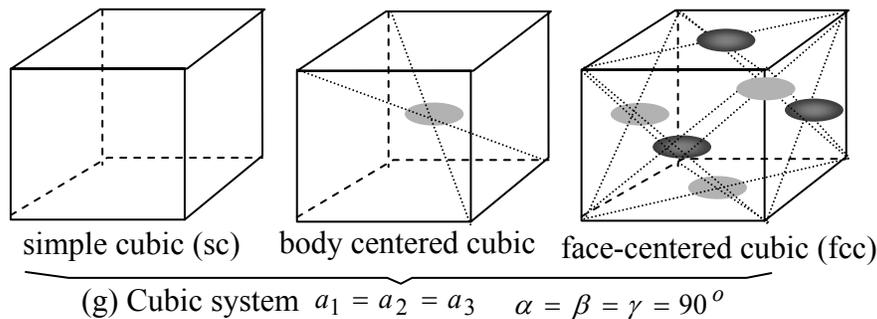


A **primitive unit cell** is the parallelepiped defined by the primitive lattice vectors. There is only one lattice point (equivalently speaking) per primitive unit cell. For example, each of the four lattice points in the two parallelograms formed by the two sets of primitive lattice vectors in above figure is shared by four unit cells and thus the number of equivalent lattice point in each parallelogram is one. These are thus primitive unit cells. On the other hand, the shaded rectangle formed by \mathbf{a}_1' and \mathbf{a}_2' are not a primitive unit cell because there are two lattice points in such a rectangle---the center point plus the four corners, each of the latter is shared by four cells. Because the choices of primitive lattice vectors are not unique, there can be different ways to draw a primitive unit cell, as the two examples in the figure. One method to construct a unit cell uniquely is the Wigner-Seitz cell, which is constructed by connecting all the neighboring points surrounding an arbitrary lattice point and drawing the bisecting plane perpendicular to each connection line. The smallest space formed by all the bisecting planes is a Wigner-Seitz cell, as indicated in the figure.

Sometimes, it is more convenient to describe a lattice by the **conventional unit cell**. For example, in the figure, the rectangle formed by \mathbf{a}_1' and \mathbf{a}_2' is more convenient than the parallelogram formed by the primitive lattice vectors. This unit cell has two lattice points and is called a conventional unit cell. The crystal can also be constructed by repeating such a conventional unit cell.

(2) Three important systems

There are totally 14 types of Bravais lattices and can be further grouped into 7 types of point symmetry operations (seven crystal systems). Here we will focus on the cubic system.



A simple cubic (sc) lattice totally encloses one atom point inside because the eight atoms at the corner are also shared by adjacent lattices. The body centered cubic (bcc) lattice has two atoms inside, while the face-centered cubic (fcc) lattice has four enclosed atoms.

Note: We can estimate the atom number for a cubic micrometer volume as

$$\left(\frac{1e-6}{5e-10}\right)^3 = 8e9,$$

which indicates the material should be similar to bulk crystal.

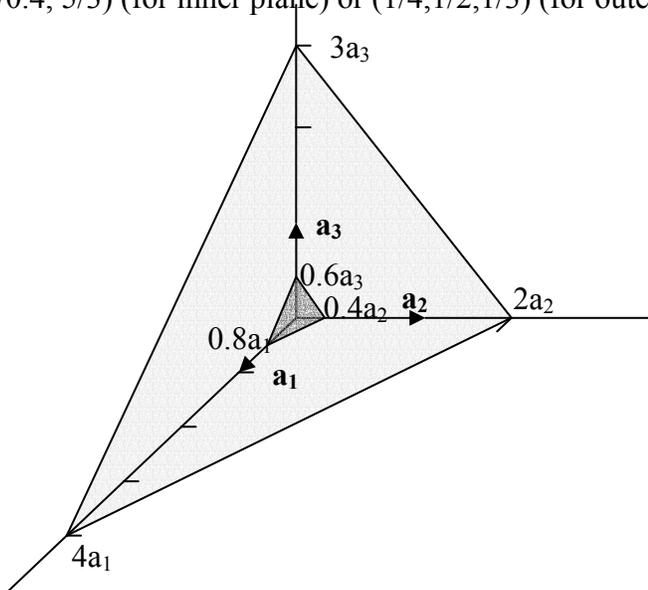
(3) Miller index

The **Miller indices of crystal planes** (hkl) are obtained in accordance with the following steps:

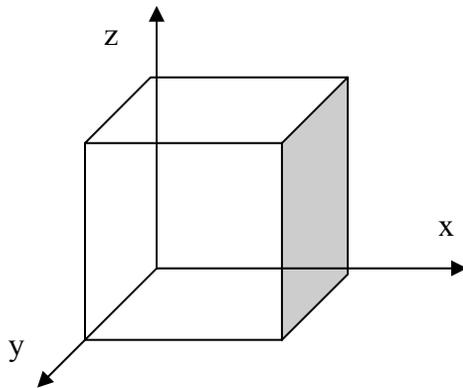
(a) Find the intercepts of the crystal plane with the axes formed by the lattice vectors \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{a}_3 in terms of the lattice constants. The origin of the lattice vectors can be at any lattice point. One can choose any crystal plane that is convenient to use. For example, in the following figure, we have two crystallographically identical planes, one intercepts the axis at $0.8a_1$, $0.4a_2$, $0.6a_3$ and the other at $4a_1$, $2a_2$, $3a_3$.

(b) Take the reciprocal of these numbers and reduce the numbers to the three smallest integers that have the same ratio as the original set. The result is enclosed in parenthesis (hkl) and this set of numbers is called the Miller indices of the plane. The above example yields

$(1/0.8, 1/0.4, 5/3)$ (for inner plane) or $(1/4, 1/2, 1/3)$ (for outer plane) $\rightarrow (364)$.

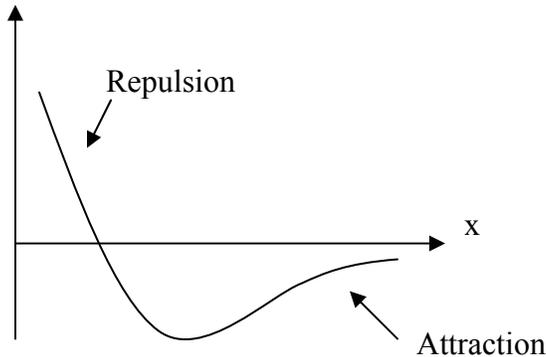


Note: Taking the reciprocal is necessary because a plane parallel to an axis will intersect it at infinity. This manipulation avoids infinity in the expression. For instance, the shaded plane in the following figure is denoted by (100). A parallel plane intersecting at $x = -1$ is denoted by $(\bar{1}00)$. We can use the sign $\{100\}$ to denote all the equivalent planes.



3.1.3 Bonding potential

Interatomic
potential Φ



Recall the previous discussion about the interatomic forces. Here we will learn them in more details.

The force interaction between atoms consists of a long-range attractive force and a short-range repulsive force. The short-range repulsive force is effective when the inner-shell electrons or the nuclei begin to overlap, due to the Pauli exclusion principle. Two often-used empirical expressions for the repulsive potential between the atoms separated by a distance r are

$$U_R(r) = \frac{B}{r^{12}} \quad (\text{Lennard-Jones})$$

and

$$U_R(r) = U_0 e^{-r/\zeta} \quad (\text{Born-Mayer})$$

where B , ζ , and U_0 are empirical constants determined from experimental data, such as the interatomic spacing and the binding energy.

Note: The repulsive forces are normally very strong and the curve is sharp when the distance approaches zero.

Molecular crystals are characterized by the dipole-dipole interaction between atoms. An isolated atom is not polarized, but when another atom is close by, the electrical field of electrons from the neighboring atom distorts the positions of the electrons and the nucleus of the current atom, creating an induced dipole. The attractive potential between the induced dipoles of two atoms is given by

$$U_A = -\frac{A}{r^6}.$$

Combining this attractive potential (van der Waals potential) with the Lennard-Jones potential for the repulsive force, we obtain the Lennard-Jones interaction potential between a pair of atoms i and j in a crystal as

$$U_{ij} = \frac{B}{r_{ij}^{12}} - \frac{A}{r_{ij}^6}.$$

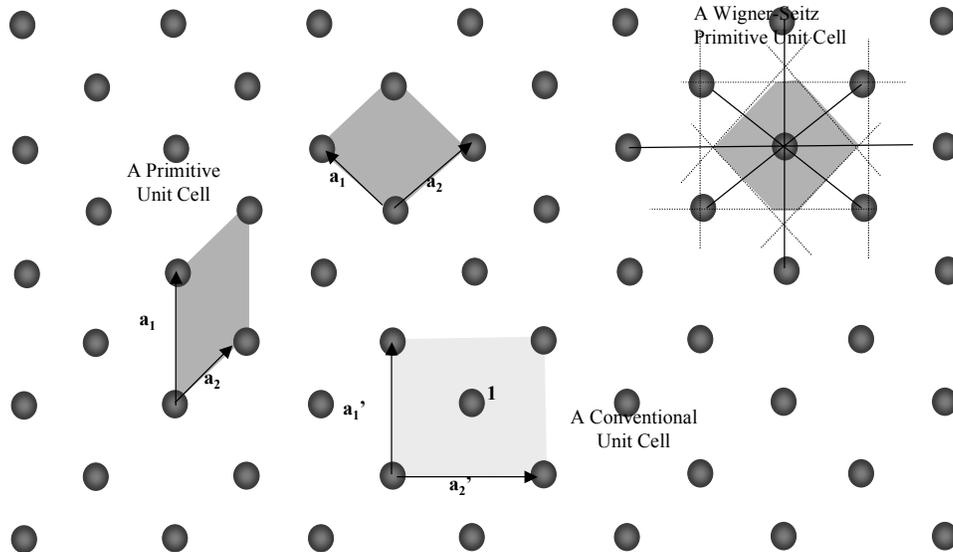
What makes a crystal structure a favorable structure is that the total potential energy of the system

$$U = \frac{1}{2} \sum_{i \neq j} \left(\frac{B}{r_{ij}^{12}} - \frac{A}{r_{ij}^6} \right) = \frac{1}{2} \sum_{i \neq j} 4\epsilon \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right)$$

reaches a minimum, as required by the second law of thermodynamics for a stable system.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 8

In the last lecture, we have talked about the **primitive unit cell**. There is only one lattice point (equivalently speaking) per primitive unit cell. The smallest space formed by all the bisecting planes is a Wigner-Seitz cell, as indicated in the figure.



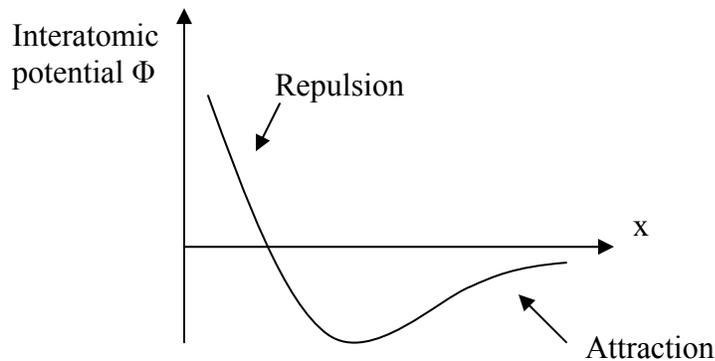
For the bonding potential, two often-used empirical expressions for the repulsive potential between the atoms separated by a distance r are

$$U_R(r) = \frac{B}{r^{12}} \quad (\text{Lennard-Jones})$$

and

$$U_R(r) = U_0 e^{-r/\zeta} \quad (\text{Born-Mayer})$$

where B , ζ , and U_0 are empirical constants determined from experimental data, such as the interatomic spacing and the binding energy.



Combining this attractive potential (van der Waals potential) with the Lennard-Jones potential for the repulsive force, we obtain the Lennard-Jones interaction potential between a pair of atoms i and j in a crystal as

$$U_{ij} = \frac{B}{r_{ij}^{12}} - \frac{A}{r_{ij}^6}.$$

What makes a crystal structure a favorable structure is that the total potential energy of the system reaches a minimum, as required by the second law of thermodynamics for a stable system.

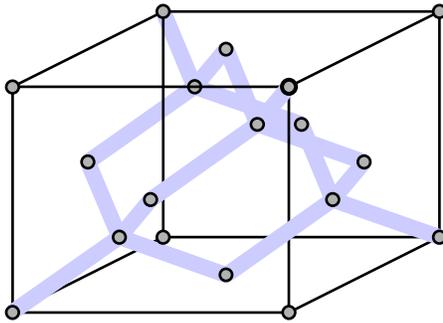
In **ionic crystals**, such as NaCl, the single valence electron in the sodium atom moves to the chlorine atom such that both Na^+ and Cl^- have closed electron-shells but meanwhile, become charged. The Coulomb potential among the ions becomes the major attractive force. The potential energy of any ion i in the presence of other ions j is then,

$$U_{i,A} = \sum_{i \neq j} \frac{\pm q^2}{4\pi\epsilon_0 r_{ij}} = -\frac{\alpha q^2}{4\pi\epsilon_0 r_0}$$

where q is the charge per ion, ϵ_0 the dielectric constant, and r_0 the nearest-neighbor separation. The parameter α is called the Madelung constant and is related to the crystal structure. This attractive potential, combined with an appropriate repulsive potential, gives a description of the potentials for ionic crystals.

Covalent bonds are formed when electrons from neighboring atoms share common orbits, rather than being attached to individual ions as in ionic crystals. Diamond, silicon, and germanium are all **covalent crystals**. Each atom has four electrons in the outer shell and forms a tetrahedral system of covalent bonds with four neighboring atoms, as indicated in the figure.

Note: For instance, in silicon the $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ atom is shared by neighboring atoms.



In certain crystals, such as GaAs, both the covalent and ionic bonding are important. Fundamentally, the covalent bonding force is still due to charge interaction. Unlike the van der Waals force in molecular crystals or the electrostatic force in ionic crystals, however, it is more difficult to write down simple expressions for the covalent crystals.

In covalent bonds, electrons are preferentially concentrated in regions connecting the nucleus, leaving some regions in the crystal with low charge concentration. Metals and their associated **metallic bonds** can be considered as an extreme case of the covalent bonds, when the covalent bonds begin to overlap and all regions of the crystal get filled up with charges. In the case of total filling of the empty space, it becomes impossible to tell which electron belongs to which atom. This can be shown by the distribution of wavefunction $\Psi(\vec{r})$, in which the probability $\Psi(\vec{r})\Psi^*(\vec{r})$ is more uniform in a metal.

3.1.4 Reciprocal lattice

If a time function $f(t)$ is periodic with a period of T (i.e. $f(t+T) = f(t)$), it can be expanded into a Fourier series as,

$$f(t) = \sum_{n=-\infty}^{\infty} \left(a_n \sin\left(\frac{2\pi n}{T}t\right) + b_n \cos\left(\frac{2\pi n}{T}t\right) \right)$$

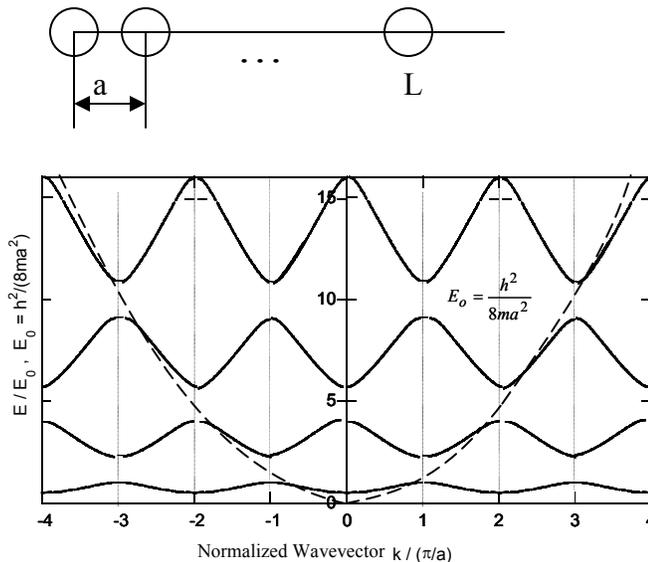
$$= \sum_{n=-\infty}^{\infty} \left(a_n 'e^{in\omega t} + b_n 'e^{-in\omega t} \right)$$

Here the angular frequency $\omega=2\pi/T$ is the Fourier conjugate of the time periodicity such that $e^{i\omega T}=1$, which ensures that $f(t)$ is periodic.

A spatial function, $f(x)$, with a periodicity a , $f(x) = f(x+a)$, can be similarly expanded into a Fourier series,

$$f(x) = \sum_{n=-\infty}^{\infty} \left(a_n 'e^{ink_x x} + b_n 'e^{-ink_x x} \right)$$

where the wavevector, $k_x=2\pi/a$, is the Fourier conjugate to spatial periodicity a .



In the above figure, the electron energy dispersion shows a period of $2\pi/a$ because of the periodic potential field $u(x+a) = u(x)$ (recall the Bloch theorem).

Note: (1) The discussed function $f(x)$ can represent not only the charge density but also other periodically distributed properties. (2) The Born-von Karman boundary condition requires that the wave functions at the two end points be equal to each other. This results in $\Delta k = \frac{2\pi}{L} = \frac{2\pi}{Na}$, not $k_x = \frac{2\pi}{a}$. (3) For personal interests, you may want to compare the process with digital signal processing (DSP), where a is in analog to the sampling period, k corresponds to the detected frequency. The highest measurable frequency is inversely proportional to a .

Now we will extend our discussion to three-dimensional cases. With a translation lattice vector $\vec{T} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$, we can reach any crystal point from the origin.

For a function $u(\mathbf{r}) = u(\mathbf{r} + \mathbf{T})$, we will first give the following answer and then show that the given Fourier expansion indeed satisfies the required periodicity,

$$u(\mathbf{r}) = \sum_{\mathbf{G}} u_{\mathbf{G}} e^{i\mathbf{r} \cdot \mathbf{G}}$$

where \mathbf{G} and the inverse transformation are given by

$$\mathbf{G} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3$$

and $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ are conjugated to the primitive lattice vectors $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ through

$$\mathbf{b}_1 = 2\pi(\mathbf{a}_2 \times \mathbf{a}_3) / V, \quad \mathbf{b}_2 = 2\pi(\mathbf{a}_3 \times \mathbf{a}_1) / V, \quad \mathbf{b}_3 = 2\pi(\mathbf{a}_1 \times \mathbf{a}_2) / V$$

where $V = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$ is the volume of the primitive unit cell in real space. For the one-dimensional case, we have $\mathbf{G} = 2\pi n/a$, $r = x$.

With the above definitions, we can show that $u(\mathbf{r})$ is indeed invariant with any translational lattice vector in the real space, $\mathbf{T} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$, where n_1, n_2, n_3 integers,

$$\begin{aligned} u(\mathbf{r} + \mathbf{T}) &= \sum_{\mathbf{G}} u_{\mathbf{G}} e^{i(\mathbf{r} + \mathbf{T}) \cdot \mathbf{G}} = \sum_{\mathbf{G}} u_{\mathbf{G}} e^{i\mathbf{r} \cdot \mathbf{G} + i\mathbf{T} \cdot \mathbf{G}} \\ &= \sum_{\mathbf{G}} u_{\mathbf{G}} e^{i\mathbf{r} \cdot \mathbf{G} + i2\pi(n_1m_1 + n_2m_2 + n_3m_3)} = \sum_{\mathbf{G}} u_{\mathbf{G}} e^{i\mathbf{r} \cdot \mathbf{G}} = u(\mathbf{r}) \end{aligned}$$

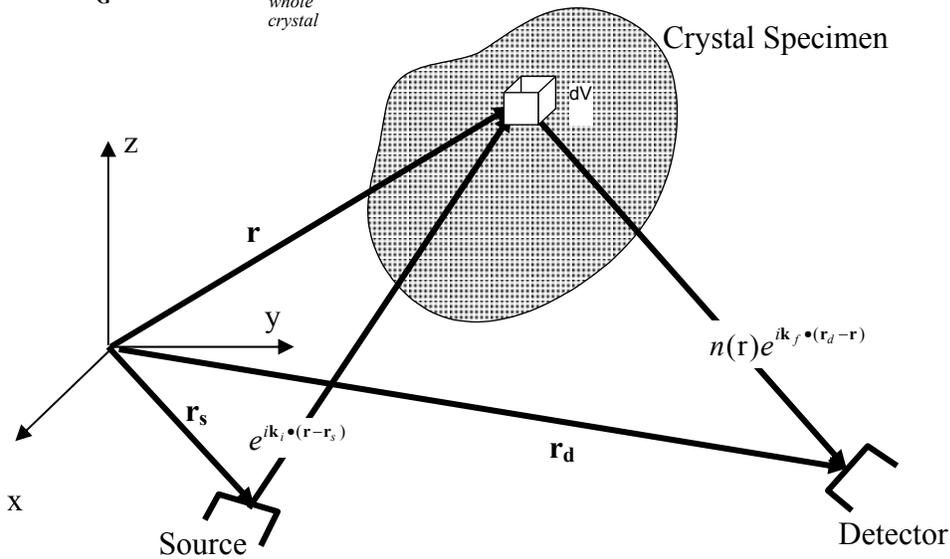
Thus, we see that the new set of vectors introduced, $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$, which has a unit of m^{-1} , are the corresponding Fourier conjugate to the real space lattice vector $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$. We can use $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ to construct a new lattice called the **reciprocal lattice**. Previous definitions on real space lattices, such as unit cells and the Wigner-Seitz primitive unit cell, are equally applicable to such a reciprocal lattice. This reciprocal space is the Fourier conjugate of the real space.

Although a very abstract concept, the reciprocal lattice can actually be easily mapped out with diffraction experiments. When electron waves or X-rays (electromagnetic waves) with proper energy are shone onto a crystal, the reflection or transmission occurs only along specific directions, as shown in the following figure.

Consider an incident wave from the source along direction \mathbf{k} . The incident wave is proportional to $e^{i\mathbf{k}_i \cdot (\mathbf{r}-\mathbf{r}_s)}$, where \mathbf{r} is any point in the sample and \mathbf{r}_s is the location of the source relative to the origin of coordinates. The wave scattered into the detector is then proportional to $n(\mathbf{r})e^{i\mathbf{k}_f \cdot (\mathbf{r}_d-\mathbf{r})}$, where \mathbf{k}_f is the propagation direction of the scattered wave, $n(\mathbf{r})$ is the nuclei density, and \mathbf{r}_d is the position of the detector. Because each of the atoms may scatter wave, the total amplitude at the detector is

$$\int_{\text{whole crystal}} A e^{i\mathbf{k}_i \cdot (\mathbf{r}-\mathbf{r}_s)} n(\mathbf{r}) e^{i\mathbf{k}_f \cdot (\mathbf{r}_d-\mathbf{r})} dV = A e^{i(\mathbf{k}_f \cdot \mathbf{r}_d - \mathbf{k}_i \cdot \mathbf{r}_s)} \int_{\text{whole crystal}} n(\mathbf{r}) e^{i(\mathbf{k}_i - \mathbf{k}_f) \cdot \mathbf{r}} dV$$

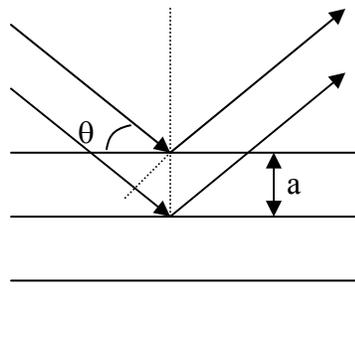
$$= A \sum_{\mathbf{G}} e^{i(\mathbf{k}_f \cdot \mathbf{r}_d - \mathbf{k}_i \cdot \mathbf{r}_s)} \int_{\text{whole crystal}} n_{\mathbf{G}} e^{i(\mathbf{G} + \mathbf{k}_i - \mathbf{k}_f) \cdot \mathbf{r}} dV$$



Because the exponential function $e^{i(\mathbf{G} + \mathbf{k}_i - \mathbf{k}_f) \cdot \mathbf{r}}$ is a rapidly varying function in the crystal with both negatives and positives, the above integral will be close to zero except when the exponent vanishes, i.e., when

$$\mathbf{G} + \mathbf{k}_i - \mathbf{k}_f = \mathbf{0}.$$

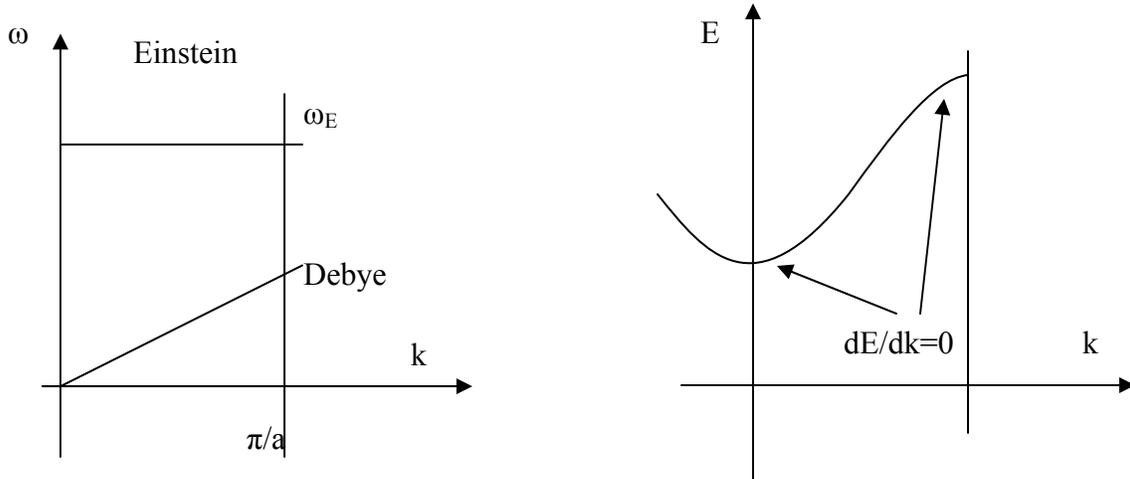
This is called the Bragg condition for diffraction. Because the wavevectors \mathbf{k}_f and \mathbf{k}_i are determined by the relative positions of the source, the sample, and the detector. The reciprocal lattice vectors \mathbf{G} , and thus the crystal structure, can be determined from diffraction experiments.



Consider the special set of crystal planes separated by a distance a as shown in the following figure and an incident wave (photon or electron) of wavelength λ at an angle θ . Constructive interference between waves reflected from crystal planes occurs when the phase difference of the waves scattered between two consecutive planes is $2\pi n$. From the figure, we see that the path difference is $2a\sin\theta$. Thus diffraction peaks will be observed when the path difference is multiples of the wavelength, i.e.,

$$2a \sin \theta = n\lambda .$$

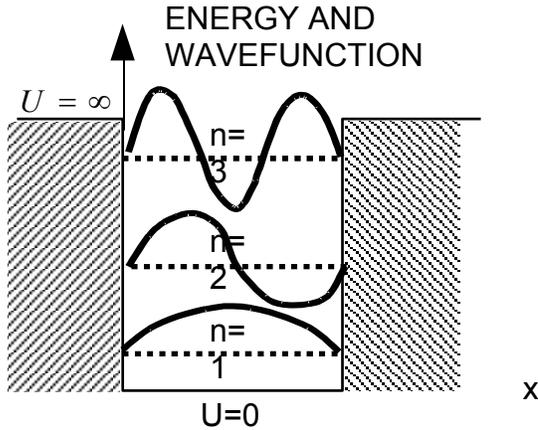
In the following lecture, we will continue the discussion of energy dispersion for electrons and phonons. For phonons, in the acoustic waves a linear energy dispersion function is used as the approximation. For electrons, two parabolic curves are connected to approximate the energy dispersion.



2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 9

3.4 Density of states

(1) Electron in a quantum well



For electrons in a quantum well, the energy has discrete levels as

$$E = \frac{\hbar^2 n^2}{8m D^2} \quad (n=1,2,\dots)$$

For wavefunction $\Psi_{n,s}$, we have degeneracy $g(n)=2$ due to the spin.

(2) Harmonic oscillator

The energy is

$$E_n = \hbar\nu(n + 1/2); \nu = \frac{1}{2\pi} \sqrt{\frac{K}{m}} \quad (n=0,1,2,\dots)$$

The wavefunction is Ψ_n , and the degeneracy is $g(n)=1$.

(3) Rigid rotation

The energy eigenvalues are

$$E_l = \frac{\hbar^2}{2I} \ell(\ell + 1) = \hbar B \ell(\ell + 1) \quad (\text{for } |m| \leq \ell, \ell = 0, 1, 2, \dots).$$

For wavefunction Ψ_{nlm} , the degeneracy is $g(l)=2l+1$.

(4) Hydrogen atom

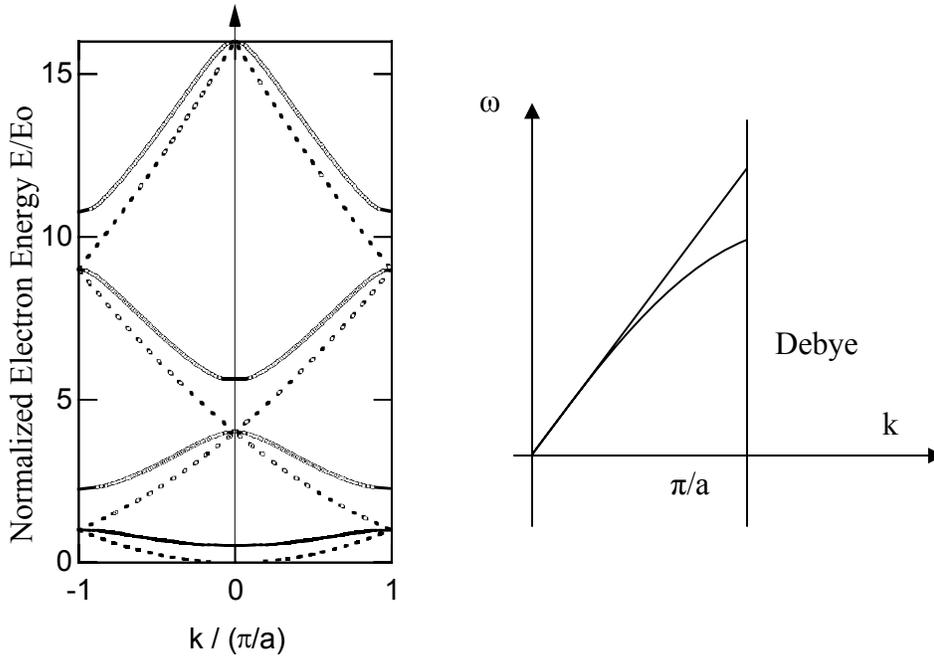
$$E_n^{el} = -\frac{Mc_1^2}{2\hbar^2 n^2} = -\frac{13.6 \text{ eV}}{n^2} \quad (n \geq 1, n \geq \ell + 1 \text{ and } |m| \leq \ell, \ell = 0, 1, 2, \dots)$$

The wavefunction Ψ_{nlms} corresponds to degeneracy $g(n)=2n^2$.

Now let us consider electrons in a solid. The parabolic approximation at the band edge gives

$$E - E_c = \frac{\hbar^2 k^2}{2m^*} = \frac{\hbar^2 (k_x^2 + k_y^2 + k_z^2)}{2m^*},$$

where $k_x, k_y, k_z = \pm \frac{2\pi na}{L}$. In different directions, n values can be different. For wavefunction $\Psi(k_x, k_y, k_z)$, we have three quantum numbers.



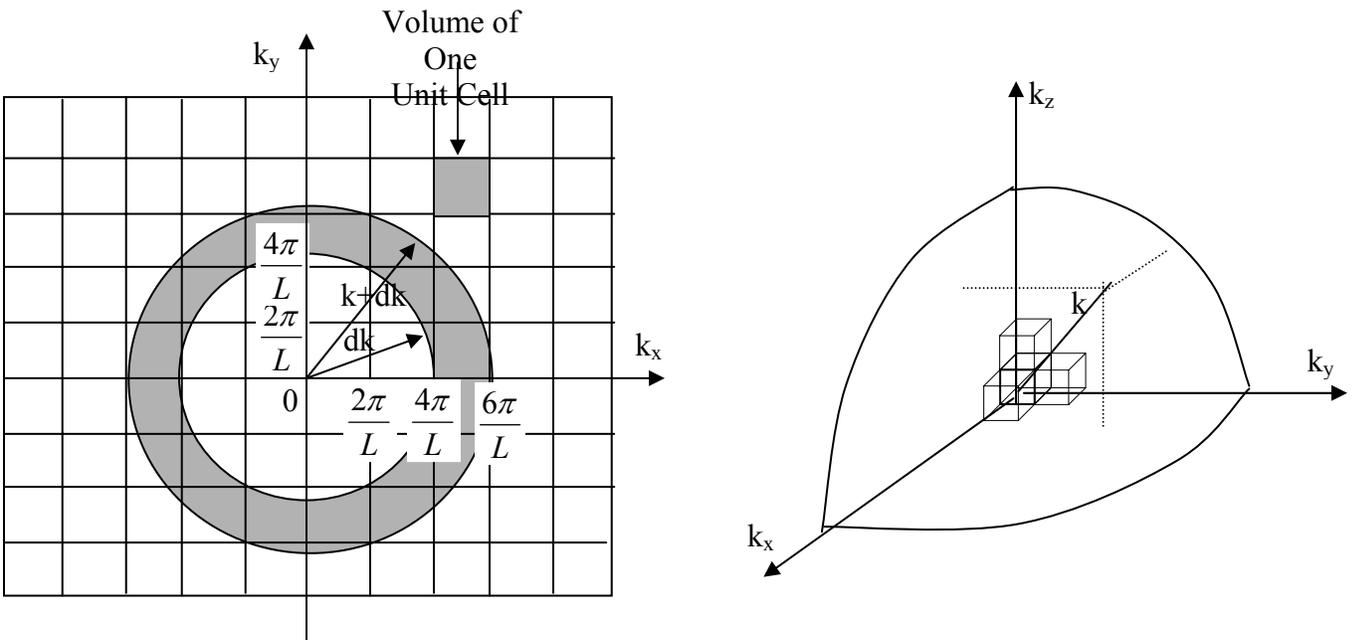
In the Debye approximation, we have energy dispersion as

$$\omega = vk = v\sqrt{k_x^2 + k_y^2 + k_z^2}; E_n = hv(n + \frac{1}{2}),$$

where v is sound velocity.

Density of (quantum mechanical) states (DOS):

(a) Electron



In above figure, we can find the volume of one state is $V_1 = (2\pi/L)^3$. In the above sphere, the number of states within k and $k+dk$ is

$$\Delta N = \frac{4\pi k^2 \Delta k}{V_1} = \frac{V k^2 \Delta k}{2\pi^2},$$

in which $V=L^3$ is the crystal volume.

The density of states is defined as the number of quantum states per unit interval of energy and per unit volume

$$D(E) = \frac{1}{V} \frac{\Delta N}{\Delta E} = \frac{k^2}{2\pi^2} \frac{\Delta k}{\Delta E} = \frac{k^2}{2\pi^2} \frac{dk}{dE}.$$

At the band edges, the electrons have energy

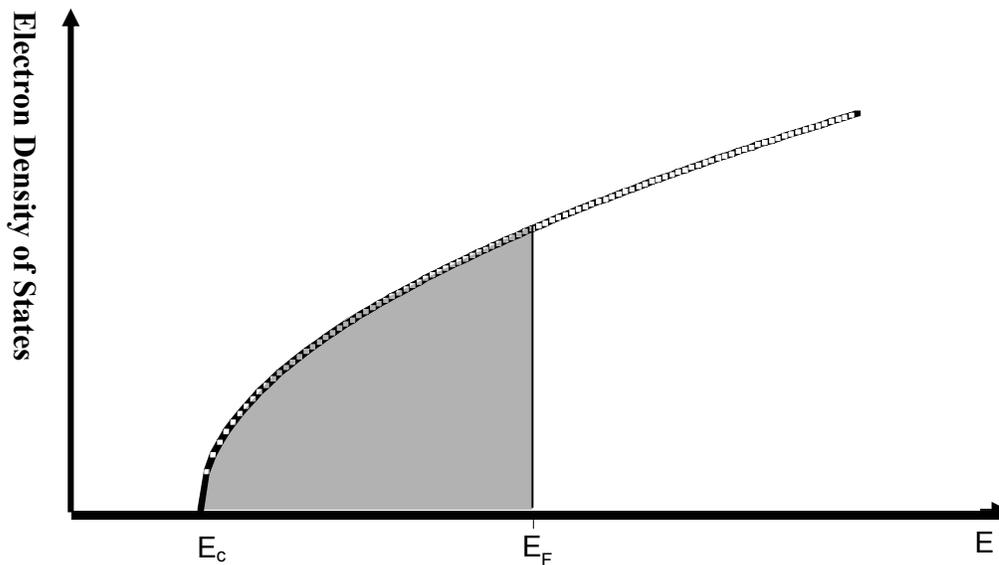
$$E = E_c + \frac{\hbar^2 k^2}{2m^*} = E_c + \frac{\hbar^2 (k_x^2 + k_y^2 + k_z^2)}{2m^*},$$

thus

$$k = \sqrt{\frac{2m^*}{\hbar^2} (E - E_c)}.$$

Considering spin, the density of states is

$$D(E) = 2 \frac{k^2}{2\pi^2} \frac{dk}{dE} = \frac{1}{2\pi^2} \left(\frac{2m^*}{\hbar^2}\right)^{3/2} (E - E_c)^{1/2}.$$



(b) Phonon

According to the Debye model that assumes three modes (two transverse, one longitudinal) are identical, we have

$$D(\omega) = 3 \frac{\Delta N}{V \Delta \omega} = \frac{3}{2\pi^2} k^2 \frac{dk}{d\omega} = \frac{3\omega^2}{2\pi^2 v^3}; \omega = vk.$$

(c) Photon

For electromagnetic waves, we only have two transverse modes. The density of states is

$$D(\omega) = \frac{\omega^2}{\pi^2 c^3}.$$

We may also want to use $D(\lambda)$ in radiation problems.

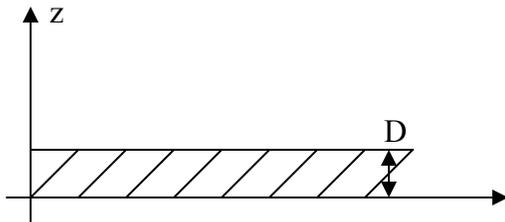
The density of states is a purely mathematical convenience but is central to correct counting of the number of electrons and the energy (or charge and momentum) that they carry. As a simple example of how the density of states is needed, let's evaluate the energy of the topmost level at $T=0$, i.e., the Fermi level, E_f . At zero Kelvin, the filling of the electron quantum states starts from the lowest energy level and moves up from one energy level to the next until all electrons are placed into distinct quantum states. The number of electrons per unit volume at $T=0$ is

$$n = \int_{E_c}^{E_f} D(E) dE = \frac{1}{3\pi^2} \left(\frac{2m^*}{\hbar^2} \right)^{3/2} (E_f - E_c)^{3/2},$$

in which the electron density $n=N/V$.

For nonzero temperatures, E_f is replaced by the chemical potential.

Nanostructures:



For a thin film with thickness d , we have the energy as

$$E(k_x, k_y, n) = \frac{\hbar^2 k_{xy}^2}{2m^*} + n^2 \frac{\hbar^2 \pi^2}{2m^* d^2} \quad (n=1,2,\dots)$$

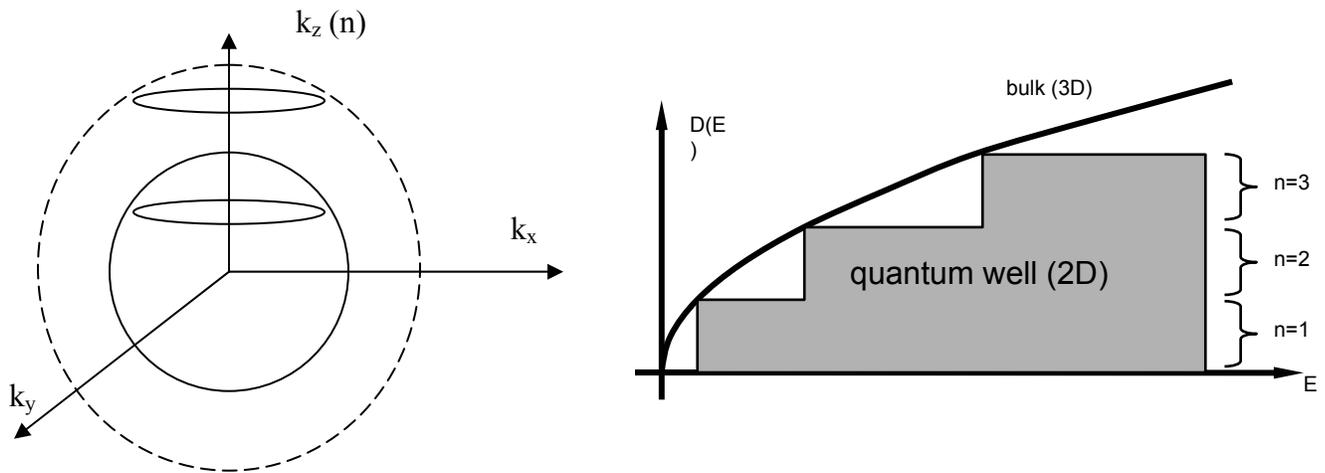
And we define $E_n = n^2 \hbar^2 \pi^2 / (2m^* d^2)$ for convenience. In the k space, $k_x, k_y = \pm \frac{2\pi ja}{L}$,

while the z direction is quantized by number n . For a given energy E , we have solutions for k_x, k_y when $E_n < E$. A constant energy surface is drawn in the following figure. Here we can nondimensionalize the coordinates to maintain the surface as a sphere. The solutions lie on the sketched circle determined by $E - E_n$, i.e.

$$|k_{xy}| = \sqrt{E(k_x, k_y, n) - E_n} \frac{\sqrt{2m^*}}{\hbar}.$$

If $E_n > E$, no solutions are available on this sphere. A large sphere corresponding to a higher E is required to find solutions for k_x, k_y .

The calculated density of states is presented in the following figure, which is staircase for one-dimensional constraint. For two-dimensional constraint (quantum wire in HW, n, l as the quantum numbers), we have $D(E) \propto (E - E_{nl})^{-1/2}$. If all three dimensions are under constraint (quantum dot, n, l, k as the quantum numbers), the density of states is just like jumps at different E_{nlk} .



Chapter 4 Statistical thermal & energy storage

In Lecture 2, we have mentioned that matter tends to occupy the lowest energy levels. For an energy level E_i , we have given the Boltzmann factor for its occupying probability as

$$P(E_i) = A \exp(-E_i / k_B T),$$

in which the constant A can be determined by normalization over all quantum states

$$\sum_{\text{All QS}} P(E_i) = 1.$$

For monatomic atom, we have

$$E = \frac{m}{2} (v_x^2 + v_y^2 + v_z^2).$$

Normalization gives

$$A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left[-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T}\right] dv_x dv_y dv_z = 1$$

$$A = \left(\frac{m}{2\pi k_B T}\right)^{3/2}.$$

The average number of particles is

$$f = \langle n \rangle = n \left(\frac{m}{2\pi k_B T} \right)^{3/2} \exp \left[-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T} \right].$$

where n is the particle number density. The Boltzmann factor can only be applied to closed systems, while for an open system exchanging energy with the outside, the probability becomes

$$P(E_i, N_i) = A \exp[-(E_i - \mu N_i) / k_B T],$$

where N_i is the particle number, chemical potential μ is the criteria for the equilibrium state of mass exchanging process with the outside, just as pressure for mechanical equilibrium and temperature for thermal equilibrium.

Now consider electrons at energy E_i . Ignoring the spin, we have two possibilities:

Electron number	N_i	Energy E_i
0	0	0
1	1	E_i

Normalization gives

$$P(0, 0) + P(E_i, 1) = 1$$

or

$$A + A \exp[-(E_i - \mu) / k_B T] = 1.$$

Therefore

$$A = \frac{1}{1 + e^{-(E_i - \mu) / k_B T}}.$$

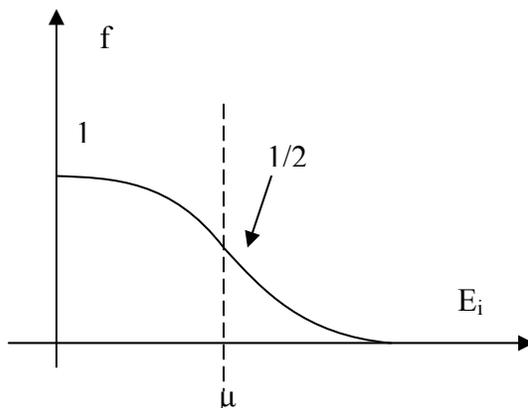
The average number of occupancy of this quantum state is thus

$$f = \langle n \rangle = 0 \times P(0, 0) + 1 \times P(E_i, 1)$$

$$= \frac{1}{1 + e^{-(E_i - \mu) / k_B T}} e^{-(E_i - \mu) / k_B T}$$

$$= \frac{1}{1 + e^{(E_i - \mu) / k_B T}}.$$

This is called the Fermi-Dirac distribution.



For phonons, the energy is expressed as

$$E_n = h\nu(n + \frac{1}{2}) \quad (n=0,1,2\dots)$$

Thus

$$\sum_{n=0}^{\infty} A \exp[-\frac{h\nu(n+1/2)}{k_B T}] = 1$$

$$\exp(-\frac{h\nu}{2k_B T}) \sum_{n=0}^{\infty} A \exp(-\frac{h\nu n}{k_B T}) = 1$$

$$\exp(-\frac{h\nu}{2k_B T}) A \frac{1}{1 - \exp(-\frac{h\nu}{k_B T})} = 1.$$

$$A = \exp(\frac{h\nu}{2k_B T}) (1 - \exp(-\frac{h\nu}{k_B T})).$$

And

$$P(E) = \exp(-\frac{nh\nu}{k_B T}) [1 - \exp(-\frac{h\nu}{k_B T})].$$

Then the average number of the phonons, or the occupancy of the quantum state is

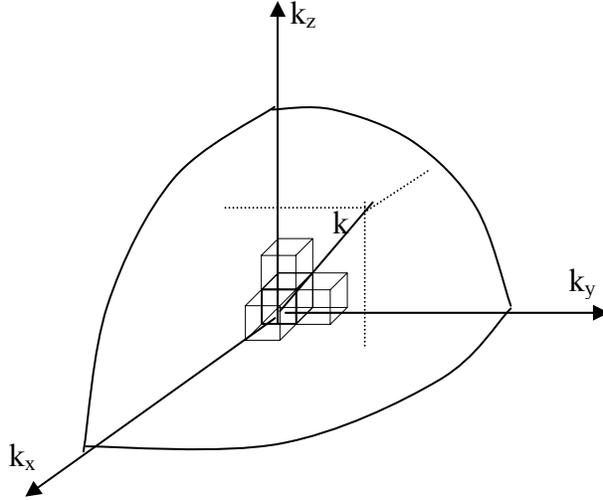
$$\begin{aligned} \langle n \rangle &= \sum_{n=0}^{\infty} n P(E_n), \\ &= [1 - \exp(-\frac{h\nu}{k_B T})] \sum_{n=0}^{\infty} n \exp(-nh\nu / k_B T) \\ &= [1 - \exp(-\frac{h\nu}{k_B T})] \left[-\frac{d}{dx} \sum_{n=0}^{\infty} \exp(-nx) \right] \quad (x=h\nu/k_B T) \\ &= \frac{1}{e^{h\nu/k_B T} - 1} \end{aligned}$$

which is the Bose-Einstein distribution.

Note: The differentiation is a mathematical trick that leads to the same results as the original summation.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 10

Review on previous lectures



In above figure, we can find the volume of one state is $V_1 = (2\pi / L)^3$. In the sphere, the number of states within k and $k+dk$ is

$$\Delta N = \frac{4\pi k^2 \Delta k}{V_1} = \frac{V k^2 \Delta k}{(2\pi)^3},$$

in which $V=L^3$ is the crystal volume.

The density of states is defined as the number of quantum states per unit interval of energy and per unit volume

$$D(E) = \frac{1}{V} \frac{\Delta N}{\Delta E} = \frac{k^2}{2\pi^2} \frac{\Delta k}{\Delta E} = \frac{k^2}{2\pi^2} \frac{dk}{dE}.$$

A factor that considers polarization of waves may be added (electron, spin up and down, thus a factor of 2, photon, two polarizations, phonons, 3 polarizations)

For an energy level E_i , the Boltzmann factor for its occupying probability is

$$P(E_i) = A \exp(-E_i / k_B T),$$

in which the constant A can be determined by normalization over all quantum states

$$\sum_{\text{All QS}} P(E_i) = 1.$$

For an open system exchanging energy with the outside, the probability becomes

$$P(E_i, N_i) = A \exp[-(E_i - \mu N_i) / k_B T],$$

where N_i is the particle number, chemical potential μ is the criteria for the equilibrium state of mass exchanging process with the outside, just as pressure for mechanical equilibrium and temperature for thermal equilibrium.

For electrons at a quantum state with energy E , the Fermi-Dirac distribution gives the average number of electrons as

$$\langle n \rangle = f(E) = \frac{1}{1 + e^{(E_i - \mu)/k_B T}}.$$

For phonons, the Pauli exclusion principle is no longer applicable. And the occupancy of the quantum state is changed into

$$\langle n \rangle = f(\nu) = \frac{1}{e^{h\nu/k_B T} - 1},$$

which is called Bose-Einstein distribution.

For molecules, similar results exist

$$\langle n \rangle = f(E) = \frac{1}{e^{(E_i - \mu)/k_B T} - 1},$$

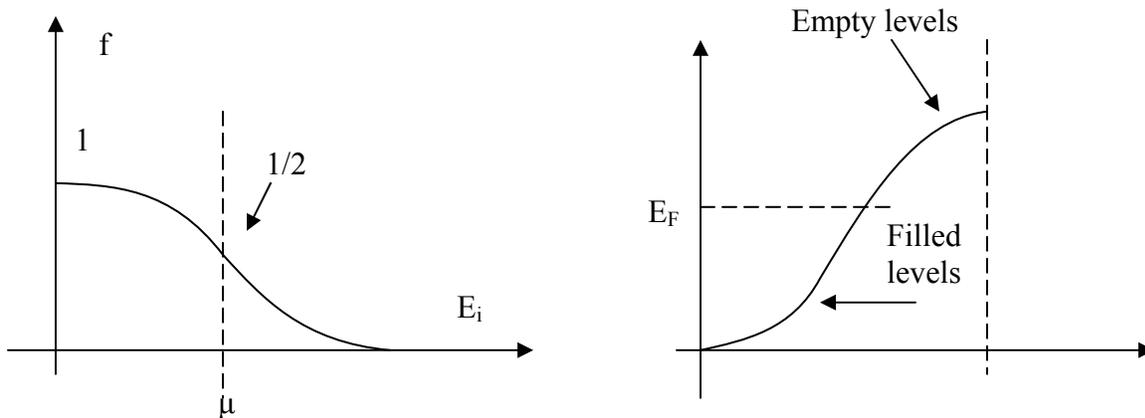
where μ is again the chemical potential of the boson gas.

The Bose-Einstein distribution changes the “plus one” in the denominator of the Fermi-Dirac distribution into minus one. When $E - \mu \gg k_B T$, we can ignore the ± 1 term in the denominator. Both distributions reduce to the Boltzmann distribution function,

$$f(E, T, \mu) = \exp\left(-\frac{E - \mu}{k_B T}\right) \text{ or } f(E, T) = \exp\left(-\frac{E}{k_B T}\right).$$

For high temperatures, the energy separation is so small that we should still use the “classic” Boltzmann distribution function.

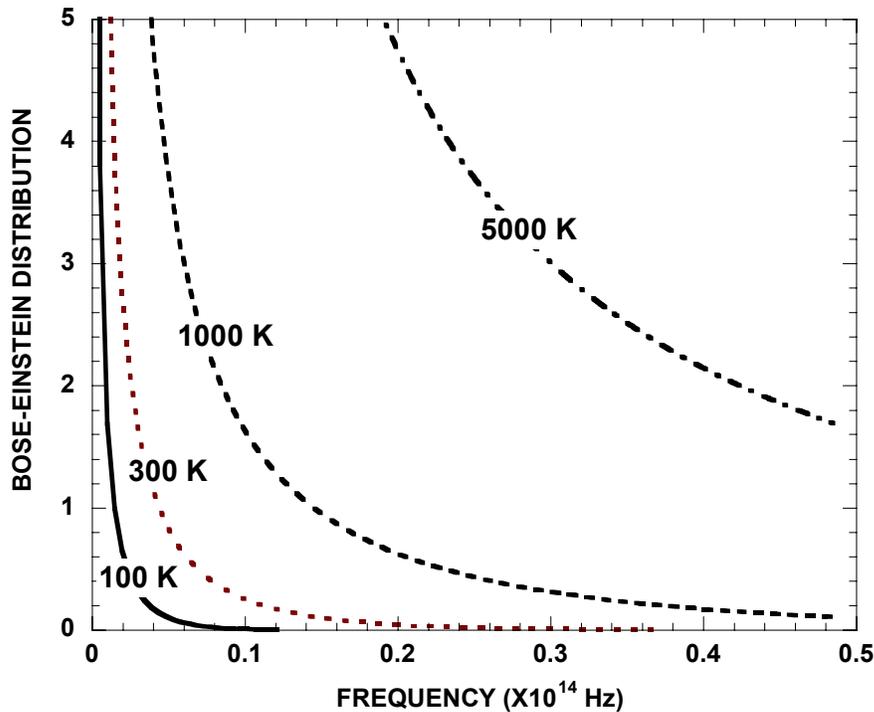
In the following figure, the Fermi-Dirac distribution is drawn. At $T=0$ K, we have $E_F = \mu$. From the right figure, we can see that electrons close to the Fermi level can be affected by the increased temperature. Only these electrons contribute to the electricity conduction.



For electrons in a box (with constraints in three dimensions), the quantized energy levels are

$$E = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2); k_x, k_y, k_z = 2\pi n / L.$$

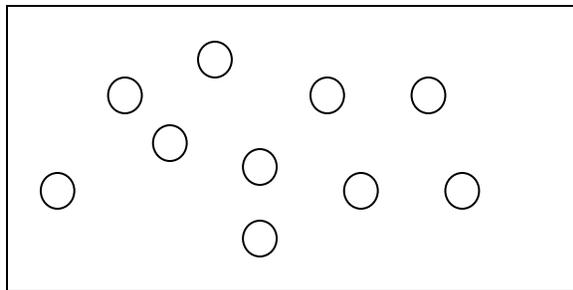
The Bose-Einstein distribution is presented in the following figure according to the temperature. For constrained electrons, we can find that the ground state (quantum numbers $n=l=i=1$) has much larger occupancy than any other energy level. This is more apparent at lower temperatures. Therefore, most molecules will go to the ground state when the temperature approaches 0 K. The phenomenon is called the Bose-Einstein condensation. The experimental work in dilute gases of alkali atoms earned Professor Ketterle a Nobel Prize in 2001.



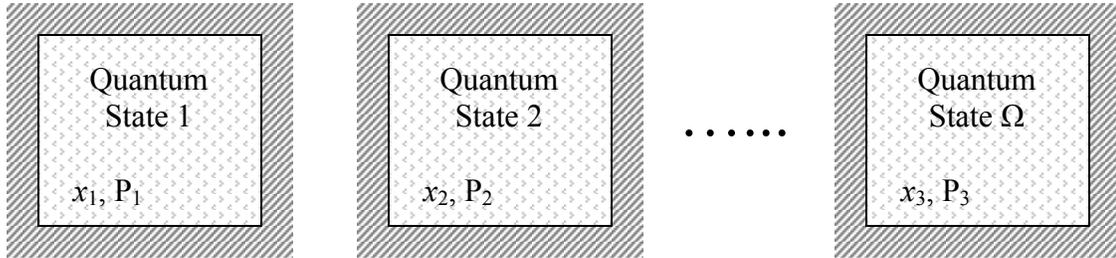
Consider particles in an isolated system. For statistics, we normally calculate the time average value

$$\bar{x} = \frac{1}{T} \int_0^T x(t) dt.$$

However, this is impractical when the particle number is very big. To simplify, we focus on the probability of a system to be at a specific accessible quantum state.



Suppose we have Ω quantum states. We can treat each accessible quantum state as a system. The collection of these Ω systems is called an ensemble. Three ensembles are analyzed in the table.



Microcanonical ensemble	Canonical ensemble	Grand canonical
<p>U, V, N fixed Isolated systems</p> <p>Principle of equal probability: $P_i = 1/\Omega$ Boltzmann principle gives $S = k_B \ln \Omega = S(U, V, N)$ The entropy S is additive.</p> <p>$dU = TdS - pdV + \mu dN$ $dS = \left(\frac{\partial S}{\partial U}\right)_{V,N} dU$ $+ \left(\frac{\partial S}{\partial V}\right)_{U,N} dV$ $+ \left(\frac{\partial S}{\partial N}\right)_{U,V} dN$</p> <p>If we know the function S(U,V,N) [or U(S,V,N)], we can determine all other thermodynamic state variables. The function S(U,V,N) is called a thermodynamic potential.</p>	<p>V, N, T fixed In contact with a thermal reservoir; isothermal.</p> <p>Since the reservoir also has many quantum states, a quantum state of one system can correspond to different number of real quantum states. The principle of equal probability is no longer valid. The probability becomes</p> <p>$P(E_i) = \frac{e^{-E_i/k_B T}}{Z}$ $Z = \sum_i e^{-E_i/k_B T}$</p> <p>The Helmholtz free energy $F = U - TS$ $= F(T, V, N)$ $= -k_B T \ln Z$</p> <p>becomes the thermal potential in this case.</p>	<p>V, μ, T fixed Open, isothermal system. Exchanging both energy and particles with the reservoir.</p> <p>$P(E_i, N_i) = \frac{e^{-(E_i - \mu N_i)/k_B T}}{Z}$</p> <p>The numerator is the Gibbs factor and $Z = \sum_{E_i} \sum_{N_i} e^{-(E_i - \mu N_i)/k_B T}$.</p> <p>The grand canonical potential is $G(T, V, \mu)$ $= U - TS - \mu N$ $= -k_B T \ln Z$</p> <p>Note: G is not Gibbs energy.</p>

Denote P as the probability of a quantum state. A fundamental assumption made in statistical mechanics is that the ensemble average of an observed quantity is equal to the time average of the same quantity, i.e.

$$\bar{x} = \langle x \rangle = \sum_{i=1}^{\Omega} P_i x_i,$$

which is called the **ergodic hypothesis**.

Internal energy and specific heat

For a constrained particle, its energy is expressed as

$$E = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2),$$

where $k_x = \pm 2\pi i / L, k_y = \pm 2\pi j / L, k_z = \pm 2\pi n / L$ ($n, j, l=1,2,\dots$). The internal energy is

$$\begin{aligned} u_1 &= \langle E_1 \rangle \\ &= \frac{\sum_{k_x} \sum_{k_y} \sum_{k_z} E(k_x, k_y, k_z) \exp\left(-\frac{E - \mu}{k_B T}\right)}{\sum_{k_x} \sum_{k_y} \sum_{k_z} \exp\left(-\frac{E}{k_B T}\right)}, \end{aligned}$$

If the energy separation is very small (quasi-continuous), we can evaluate Z by integration instead of discrete summation, i.e.

$$\begin{aligned} Z &= \sum_{k_x} \sum_{k_y} \sum_{k_z} \exp\left(-\frac{E}{k_B T}\right) \\ &= \int_0^{\infty} e^{-E/k_B T} D(E) dE \\ &= \int_0^{\infty} e^{-E/k_B T} \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \sqrt{E} dE \\ &= \frac{V}{\lambda^3}, \end{aligned}$$

where the thermal de Broglie wavelength is $\lambda = \sqrt{\frac{\hbar^2}{2\pi m k_B T}}$.

Let $y = \frac{1}{k_B T}$. The energy expression can be rewritten as

$$\langle E_1 \rangle = \frac{1}{Z} \frac{\partial Z}{\partial y} = -\frac{d}{dy} \ln Z = \frac{3}{2} k_B T,$$

which is just the familiar energy expression. It is also a special case of the **equipartition theorem**, which states that at high temperature every degree of freedom with a quadratic energy term contributes $k_B T/2$ to the average energy of the system.

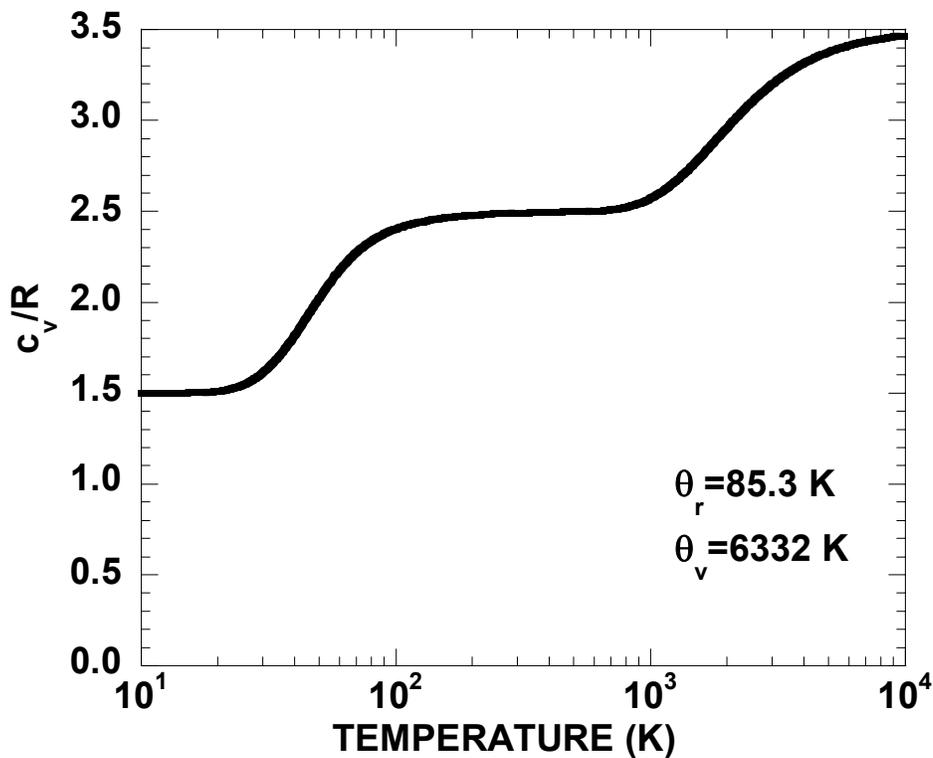
The internal energy is

$$u = \frac{3}{2} k_B N_A T = \frac{3}{2} R_u T,$$

in which R_u is the universal gas constant. The specific heat is

$$C_V = \frac{\partial u}{\partial T} = \frac{3}{2} R_u.$$

In the following figure we draw the specific heat of hydrogen gas. At low temperatures, only the translational energy levels are fully excited and the specific heat is $3R_u/2$. As the temperature increases, the rotational energy levels start to be excited and contribute to the specific heat to a maximum of R such that the total specific heat reaches $5R_u/2$. At even higher temperatures, the vibrational energy levels start contributing to the specific heat that approaches a final value of $7R_u/2$.



For photons, we have

$$\langle n \rangle = f(\omega) = \frac{1}{e^{\hbar\omega/k_B T} - 1}.$$

The internal energy is

$$u = \sum_{k_x} \sum_{k_y} \sum_{k_z} \hbar\omega f(\omega) = \int_0^\infty \hbar\omega f(\omega) D(\omega) d\omega,$$

where the density of states (two transverse electromagnetic waves) is

$$D(\omega) = 2 \times \frac{4\pi k^2 dk}{\left(\frac{2\pi}{L}\right)^3 d\omega} = \frac{\omega^2}{\pi^2 c^3}; \omega = ck \text{ for light.}$$

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 11

4.1.1 Photons (continue)

First let us continue the discussion of photons. We have

$$\langle n \rangle = f(\omega, T) = \frac{1}{e^{\hbar\omega/k_B T} - 1}.$$

The internal energy is

$$u = \sum_{k_x} \sum_{k_y} \sum_{k_z} \hbar\omega f(\omega, T) = \int_0^\infty \hbar\omega f(\omega, T) D(\omega) d\omega,$$

where the density of states (two transverse electromagnetic waves) is

$$D(\omega) = 2 \times \frac{4\pi k^2 dk}{\left(\frac{2\pi}{L}\right)^3 d\omega} = \frac{\omega^2}{\pi^2 c^3}; \omega = ck \text{ for light.}$$

Thus

$$u = \int_0^\infty \hbar\omega \frac{\omega^2}{\pi^2 c^3} \frac{1}{[\exp(\hbar\omega/k_B T) - 1]} d\omega = \int_0^\infty u_\omega d\omega,$$

in which we define

$$u_\omega = f(\omega, T) \hbar\omega D(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{[\exp(\hbar\omega/k_B T) - 1]}.$$

Recall the Planck's law as

$$e_{b,\lambda} = \frac{c_1}{\lambda^5 (e^{c_2/\lambda T} - 1)},$$

which has λ^5 but in u_ω we have ω^3 . Noticing $u_\lambda = \left| \frac{\Delta\omega}{\Delta\lambda} \right| u_\omega$ and $\omega = ck = 2\pi c / \lambda$, we

obtain $u_\lambda = \left| \frac{d\omega}{d\lambda} \right| u_\omega = \frac{2\pi c}{\lambda^2} u_\omega$, which accounts for the power difference in above two

expressions. Therefore, we have obtained the exact Planck's blackbody radiation law. The radiation intensity, defined as energy flux per unit solid angle and normal area, is calculated as

$$I_\lambda = \frac{c u_\lambda}{4\pi},$$

where 4π is the full solid angle (for sphere $\Omega = Area / R^2 = 4\pi$), c is light speed.

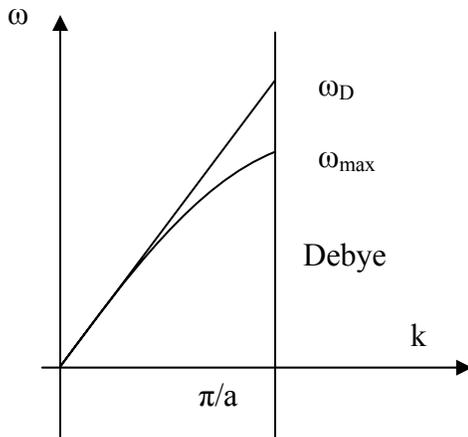
With $u = \int_0^\infty u(\omega) d\omega = \frac{4\sigma}{c} T^4$, the blackbody emissive power is expressed as

$$e_b = \pi I = \sigma T^4, \text{ where } \sigma = 5.67 \times 10^{-8} W / m^2 \cdot K^4.$$

4.1.2 Phonons

In the Debye model, we assume $\omega = v_D |k|$. The maximum frequency is thus $\omega_D = v_D \frac{\pi}{a}$, which is different from original ω_{\max} . For phonons, we have three polarizations (two transverse, one longitudinal waves). Similar to photons, we have

$$D(\omega) = \frac{dN}{V d\omega} = 3 \cdot \frac{\omega^2}{2\pi^2 v_D^3}.$$



The internal energy is

$$\begin{aligned} U &= \sum_{k_x} \sum_{k_y} \sum_{k_z} \hbar \omega f(T, \omega) \\ &= \int_0^{\omega_D} \hbar \omega f(T, \omega) D(\omega) d\omega \\ &= \frac{3}{2\pi^2 v_D^3} \int_0^{\omega_D} \frac{\hbar \omega^3 d\omega}{\exp(\hbar \omega / \kappa_B T) - 1}. \end{aligned}$$

Note: For one monoatomic chain with N atoms, we have N quantum states. Similarly, for N atoms in a crystal, we have $3N$ quantum states (three acoustic branches) for phonons. This yields

$$3N/V = \frac{1}{V} \sum_{k_x} \sum_{k_y} \sum_{k_z} 1 = \int_0^{\omega_D} D(\omega) d\omega = \int_0^{\omega_D} \frac{3\omega^2}{2\pi^2 v_D^3} d\omega = \frac{\omega_D^3}{2\pi^2 v_D^3},$$

from which we can estimate ω_D and the effective lattice constant $a = v_D \frac{\pi}{\omega_D}$.

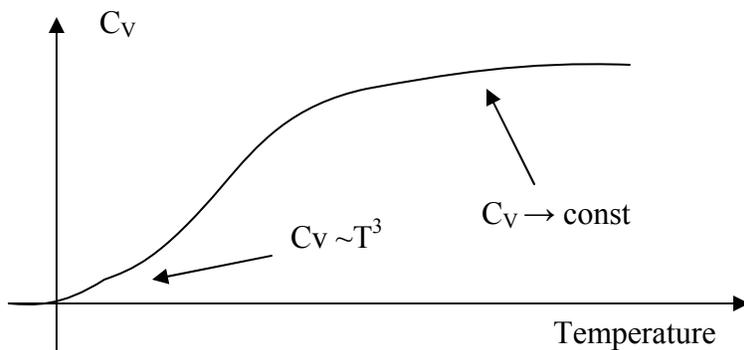
Define Debye temperature $\theta_D = \hbar \omega / k_B$. The specific heat is expressed as

$$C_V = 9k_B \frac{N}{V} \left(\frac{T}{\theta_D} \right)^3 \int_0^{\theta_D/T} \frac{x^4 e^x dx}{(e^x - 1)^2},$$

in which $x = \frac{\hbar\omega}{k_B T}$.

At low temperatures, the upper limit of above integration $\theta_D / T \rightarrow \infty$ and thus $C_V \propto T^3$, while in the high-temperature limit we have constant $C_V = 3 \frac{N}{V} k_B$.

Note: At high temperatures, in each direction both the harmonic component $\frac{Kx^2}{2}$ and kinetic component $\frac{mv^2}{2}$ contribute $k_B T / 2$ to the internal energy. Vibrations of one ion in three directions totally contribute $k_B T$, which is consistent with $C_V = 3 \frac{N}{V} k_B$. This result is similar to the ideal gas case. A common C_V - T curve is drawn as following.



In many sources, people use the specific heat per unit mass instead of per unit volume. This normally causes a factor difference. In general, we have $C_V \sim 10^6 \text{ J / K} \cdot \text{m}^3$.

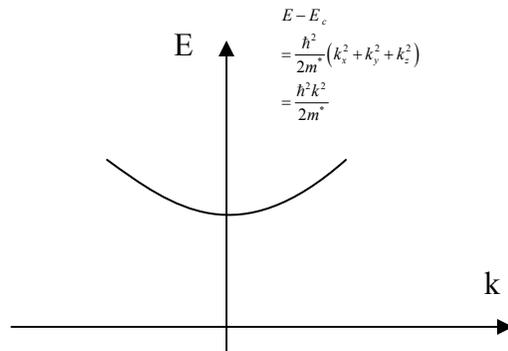
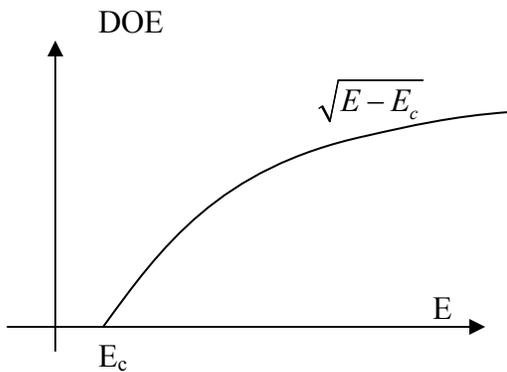
4.1.3 Electrons

In this case, the internal energy per unit volume is

$$U(T) = 2 \sum_{k_x} \sum_{k_y} \sum_{k_z} E f(E, T, \mu) = \int_{E_c}^{\infty} E f(E, T, \mu) D(E) dE,$$

where the Fermi-Dirac distribution is $f = \frac{1}{e^{(E-\mu)/k_B T} + 1}$, the coefficient 2 for spins is

included in $D(E)$, the energy dispersion $E - E_c = \frac{\hbar^2}{2m^*} (k_x^2 + k_y^2 + k_z^2)$.



We can use the same idea as phonons to solve μ ,

$$\begin{aligned} \frac{N}{V} &= \frac{2}{V} \sum_{k_x} \sum_{k_y} \sum_{k_z} f(E, \mu, T) \\ &= \frac{2}{V} \iiint_V f(E, \mu, T) \left(\frac{dk_x}{2\pi/L} \right) \left(\frac{dk_y}{2\pi/L} \right) \left(\frac{dk_z}{2\pi/L} \right) \\ &= \int_0^\infty f(E, \mu, T) D(E) dE, \end{aligned}$$

where N is the total number of electrons, and we use density of state to rewrite the summation in an integral form. The chemical potential μ is solved if N/V is given.

The specific heat is derived as

$$C_e = \frac{\partial u}{\partial T} = \frac{1}{2} \pi^2 \frac{N}{V} k_B \frac{T}{T_f},$$

in which the Fermi temperature is defined as $T_f = \frac{E_f}{k_B}$.

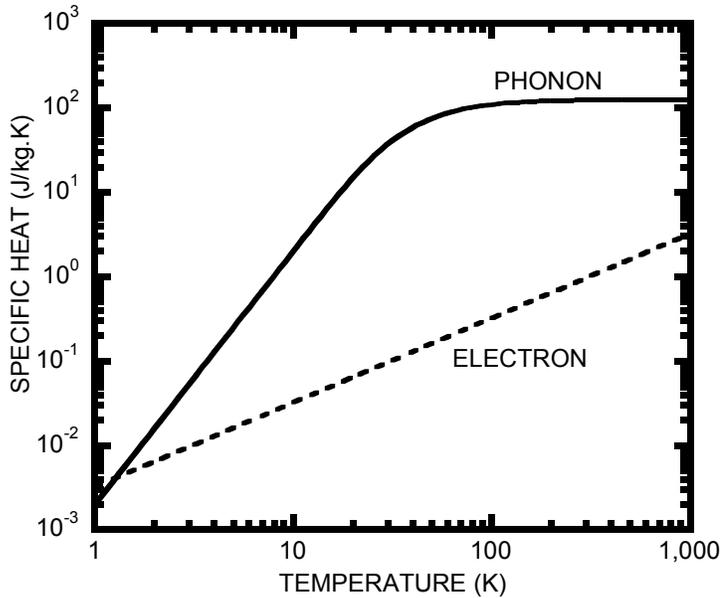
Note: (1) When we calculate $\sum_{k_x} \sum_{k_y} \sum_{k_z} f(E, \mu, T)$, the energy values within the energy

gap should be excluded and we should not count the corresponding wavevectors. (2) In a semiconductor, the phonons contribute much more to the specific heat than electrons. Electron contribution can only exceed that of the phonon at very low temperatures in the

following figure. In the equation $k = \frac{Cv\Lambda}{3}$, there are different specific heat C and

velocity v for electrons and phonons. We should also have different mean free paths. (3)

For nanostructures, the energy is quantized and the summation should be conducted over the quantum numbers instead of wavevectors, such as quantum dots. (4) All current discussions are based on equilibrium state. We will talk about nonequilibrium problems later.

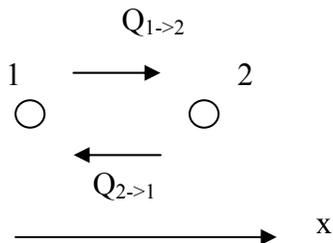


Chapter 5 Energy transport by waves

Consider energy transported between two points. The net energy transfer rate is

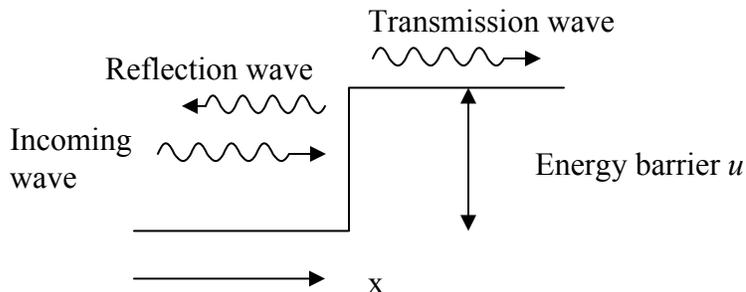
$$q_{12} = q_{1 \rightarrow 2} - q_{2 \rightarrow 1} = \int \tau_{1 \rightarrow 2} v_x E_1 f(T, E_1) D(E_1) dE_1,$$

in which $\tau_{1 \rightarrow 2}$ is the transmissivity, $E_1 f(T, E_1) D(E_1)$ has the unit J/m^3 , and q_{12} has the unit W/m^2 .



Now the question is how to calculate the transmissivity τ and velocity v . For nanostructures, the interface characteristic length is comparable to wavelength so that the interface is important even in the classic viewpoint.

5.1 Plane waves & their interface reflection



Recall the above homework problem. Generally we have the wavefunction of the transmitted wave as

$$\Psi_t(t, x) = e^{-i\frac{E}{\hbar}t} \Psi(x) = e^{-i(\omega t - k_x x)},$$

in which $k_x = \sqrt{\frac{2m(E - u)}{\hbar^2}}$.

Based on the Schrödinger equation

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi + (U - E)\Psi = 0,$$

we have

$$\Psi_i = Ae^{-i(\omega t - k_1 x)} \quad (\text{incoming wave}),$$

$$\Psi_r = Be^{-i(\omega t + k_1 x)} \quad (\text{reflected wave}), \quad k_1 = \sqrt{\frac{2mE}{\hbar^2}}$$

$$\Psi_t = Ce^{-i(\omega t - k_2 x)} \quad (\text{transmitted wave}), \quad k_2 = \sqrt{\frac{2m(E - u)}{\hbar^2}}.$$

The boundary conditions are applied

$$(\Psi_i + \Psi_r)_{x=0^-} = \Psi_t|_{x=0^+}, \quad (\Psi_i' + \Psi_r')_{x=0^-} = \Psi_t'|_{x=0^+},$$

which yields

$$A + B = C; k_1(A - B) = k_2 C$$

$$\text{or } \frac{B}{A} = \frac{k_1 - k_2}{k_1 + k_2}; \frac{C}{A} = \frac{2k_1}{k_1 + k_2}.$$

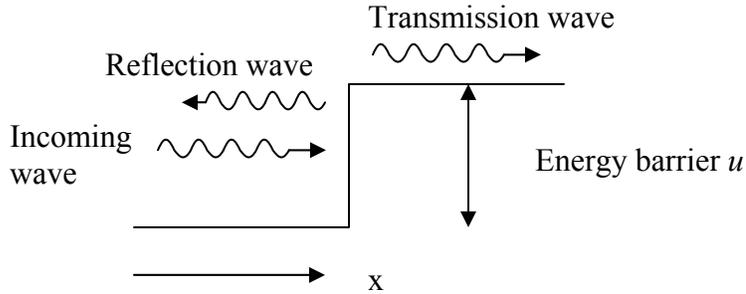
The flux term is (note A and k_1 can be complex)

$$\begin{aligned} J &= \text{Re} \left(\frac{i\hbar}{m} \Psi \nabla \Psi^* \right) \\ &= \text{Re} \left(\frac{i\hbar}{m} A e^{-i(\omega t - k_1 x)} A^* (-ik_1^*) e^{i(\omega t - k_1^* x)} \Big|_{x=0} \right) \\ &= \frac{\hbar}{m} |A|^2 \text{Re} \left(k_1^* e^{i(k_1 - k_1^*)x} \Big|_{x=0} \right) \\ &= \frac{\hbar}{m} |A|^2 k_1, \end{aligned}$$

where we use the fact that k_1 is a real number in the last step.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 12

5.1 Plane waves & their interface reflection (continue)



For the above problem, we have obtained

$$\Psi_i = Ae^{-i(\omega t - k_1 x)} \quad (\text{incoming wave}),$$

$$\Psi_r = Be^{-i(\omega t + k_1 x)} \quad (\text{reflected wave}), \quad k_1 = \sqrt{\frac{2mE}{\hbar^2}}$$

$$\Psi_t = Ce^{-i(\omega t - k_2 x)} \quad (\text{transmitted wave}), \quad k_2 = \sqrt{\frac{2m(E - u)}{\hbar^2}}.$$

The boundary conditions are applied

$$(\Psi_i + \Psi_r)_{x=0^-} = \Psi_t|_{x=0^+}, \quad (\Psi_i' + \Psi_r')_{x=0^-} = \Psi_t'|_{x=0^+},$$

which yields

$$A + B = C; \quad k_1(A - B) = k_2 C$$

$$\text{or } \frac{B}{A} = \frac{k_1 - k_2}{k_1 + k_2}, \quad \frac{C}{A} = \frac{2k_1}{k_1 + k_2}.$$

The incoming flux term is (note A and k_1 can be complex)

$$\begin{aligned} J_i &= \text{Re} \left(\frac{i\hbar}{m} \Psi_i \nabla \Psi_i^* \right) \\ &= \text{Re} \left(\frac{i\hbar}{m} A e^{-i(\omega t - k_1 x)} A^* (-ik_1^*) e^{i(\omega t - k_1^* x)} \Big|_{x=0} \right) \\ &= \frac{\hbar}{m} |A|^2 \text{Re} \left(k_1^* e^{i(k_1 - k_1^*)x} \Big|_{x=0} \right) \\ &= \frac{\hbar}{m} |A|^2 \text{Re} (k_1^*) \\ &= \frac{\hbar}{m} |A|^2 k_1, \end{aligned}$$

where we use the fact that k_1 is a real number in the last step. Similarly, we have

$$J_r = -\frac{\hbar}{m}|B|^2 k_1 \text{ (negative sign indicating the direction of reflection),}$$

$$J_t = \frac{\hbar}{m}|C|^2 \text{Re}(k_2^*).$$

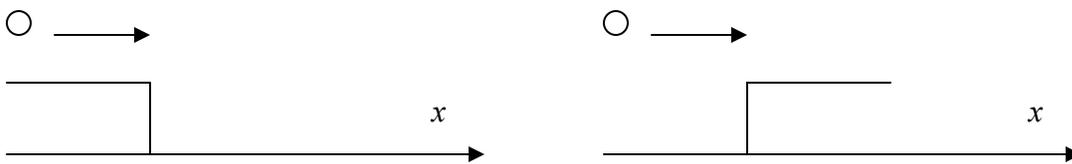
The reflectivity is

$$R = -J_r / J_i = \left| \frac{B}{A} \right|^2 = \left| \frac{k_1 - k_2}{k_1 + k_2} \right|^2 = \left| \frac{\sqrt{E} - \sqrt{E-u}}{\sqrt{E} + \sqrt{E-u}} \right|^2,$$

and transmittivity is

$$T = J_t / J_i = \left| \frac{C}{A} \right|^2 \frac{\text{Re}(k_2^*)}{k_1}.$$

For $E > u$, the equation gives reflectivity $R \neq 0$ in both cases shown below, which is inconsistent with classical mechanics.



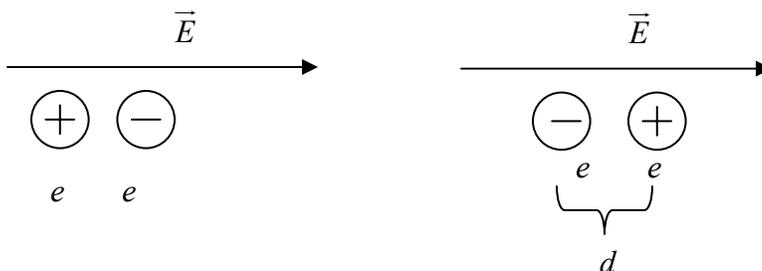
When $E < u$, the equations yield $R=1$, $T=0$, which is reasonable from classical viewpoint.

However, the wavefunction is $\Psi_t = C e^{-i\omega t - \sqrt{\frac{2m(u-E)}{\hbar^2}}x} \neq 0$ in this case. The wave will decay rapidly from the barrier and is thus called “evanescent wave.”

Electromagnetic (EM) waves

In this chapter, we will see that the wave reflection, interference, and tunneling phenomena can occur for all the three types of carriers (phonons, electrons, photons) and the descriptions of these phenomena are also similar.

An electromagnetic wave in vacuum is characterized by an **electric field vector**, \vec{E} , and a **magnetic field vector**, \vec{H} . Consider a pair of charged particles placed in an electrical field. The field will attract the positively charged particle in one direction and repel the other particle in the same direction. Consequently, the particles are distorted and form an electrical dipole, whose moment is $p = e \cdot d$ (d is the separation distance).



A measure of the capability of the material responding to incoming electric field is the electric polarization per unit volume, or the **dipole moment** per unit volume, \mathbf{P} [C m^{-2}], which is related to the electric field through the electric susceptibility, χ ,

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E},$$

where ϵ_0 is the vacuum electric permittivity, $\epsilon_0 = 8.85 \times 10^{-12}$ [$\text{C}^2 \text{N}^{-1} \text{m}^{-2}$], and the electric susceptibility is nondimensional. The electric susceptibility χ describes the extent to which positive and negative charges are displaced in a dielectric material under an applied electric field.

The total field inside the medium is measured by the **electric displacement**, \mathbf{D} [Cm^{-2}], which is a superposition of the contributions from the external electric field and the electric polarization,

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon_0 (1 + \chi) \mathbf{E} = \epsilon \mathbf{E}$$

where ϵ is called the electrical **permittivity** of the medium.

The electron and ion motion in a medium also induces a magnetic field, which is superimposed onto the external magnetic field. A measure of the total magnetic field inside the medium is called magnetic induction, \mathbf{B} ($\text{N.s m}^{-1} \cdot \text{C}^{-1}$),

$$\mathbf{B} = \mu \mathbf{H}$$

where μ is the magnetic **permeability**.

The propagation of an electromagnetic wave is governed by the following **Maxwell equations**:

$$(1) \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

This is the Farady law, which states that a changing magnetic field induces an electric field.

$$(2) \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}_e$$

Without the $\frac{\partial \mathbf{D}}{\partial t}$ term, the above equation is the Ampere law, which says an electric

current induces a magnetic field. The term $\frac{\partial \mathbf{D}}{\partial t}$ is the current due to the electron oscillation around the ion even though they are not free to move. It is also called displace current. This term is Maxwell's contribution. The current density term on the RHS is determined by $\mathbf{J}_e = \sigma \mathbf{E}$, in which σ is electrical conductivity.

$$(3) \nabla \cdot \mathbf{D} = \rho_e$$

Here ρ_e is the net charge per unit volume (C m^{-3}).

$$(4) \nabla \cdot \mathbf{B} = 0$$

It states that there is no magnetic analog of an electric charge as in (3).

Rearranging the Maxwell equations yields $\nabla^2 \mathbf{E} = \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu\sigma \frac{\partial \mathbf{E}}{\partial t}$, where the term $\mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}$ denotes displacement, while $\mu\sigma \frac{\partial \mathbf{E}}{\partial t}$ denotes the current. Without $\mu\sigma \frac{\partial \mathbf{E}}{\partial t}$, the equation $\nabla^2 \mathbf{E} = \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}$ is just a regular wave function. The additional term $\mu\sigma \frac{\partial \mathbf{E}}{\partial t}$ corresponds to damping and is also called dissipation term.

By solving equations (1)-(5), we obtain the following results for EM waves

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp[-i(\omega t - \mathbf{k} \cdot \mathbf{r})],$$

$$\mathbf{k} \cdot \mathbf{k} = \mu\omega^2 [\epsilon_0(1+\chi) + i\sigma_e/\omega] = \mu\epsilon\omega^2,$$

where the first equation has the same form as the foregoing transmitted wave. In vacuum $\sigma_e = \chi = 0$, the second equation yields

$$|\mathbf{k}|^2 = \epsilon_0 \mu_0 \omega^2,$$

or

$$\omega = \frac{|\mathbf{k}|}{\sqrt{\epsilon_0 \mu_0}} = c_0 |\mathbf{k}|,$$

in which c_0 is the light speed in vacuum. This is familiar energy dispersion relationship for photons. Compared with photons, Maxwell concluded that EM waves were the same kind of wave as light. Here we also define the **complex reflective index** N as

$$N = \frac{c_0}{c} = \sqrt{\frac{\epsilon\mu}{\epsilon_0\mu_0}} \approx \sqrt{\frac{\epsilon}{\epsilon_0}} = \sqrt{\epsilon_r} = n + ik,$$

where $\epsilon_r = \epsilon/\epsilon_0$ is called the **dielectric constant** or dielectric function, the real part of N , n , is the usual refractive index of materials. The imaginary part of N , k , is called the **extinction coefficient**, measures the damping of the electromagnetic field, which not only arises from the free electrons absorption, but also from the dipole oscillation of bounded electrons and other mechanisms.

Note: The ϵ_r value depends on the wave frequency and is not a constant.

Based on N , $|\mathbf{k}|^2 = \epsilon\mu\omega^2$ can be rewritten as

$$|\mathbf{k}|^2 = \frac{\omega^2}{c^2} = \frac{\omega^2}{c_0^2} \frac{c_0^2}{c^2} = \frac{\omega^2}{c_0^2} N^2,$$

or

$$|\mathbf{k}| = \frac{N\omega}{c_0}.$$

One can prove that the electromagnetic wave is a transverse wave, and that the electrical and magnetic fields are perpendicular to each other, i.e.,

$$\mathbf{E} \perp \mathbf{H} \perp \mathbf{k}$$

In the special case that a plane wave is traveling along the x-direction with the electric and magnetic fields pointing to the y- and z-direction, respectively.

We also define the **Poynting vector S** (W/m^2) as

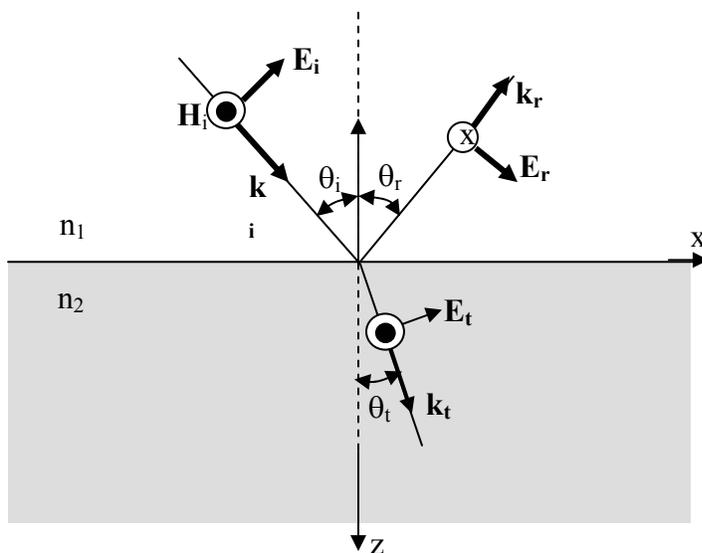
$$\mathbf{S} = \mathbf{E} \times \mathbf{H},$$

which represents the instantaneous energy flux. It oscillates at twice frequency of the electromagnetic field. No electronic devices can measure such a fast signal. What can be measured is the time-averaged Poynting vector that is further expressed as

$$\langle \mathbf{S} \rangle = \frac{1}{2} \text{Re}(\mathbf{E}_c \times \mathbf{H}_c^*)$$

Note: Normally we have $\mu, \epsilon > 0$, and thus $n > 0$ and the vector \mathbf{S} is parallel to the wavevector \mathbf{k} , which indicates the energy flows in the propagation direction of the wave. When $\mu, \epsilon < 0$, the refractive index n should take a negative value. Such media do not exist in nature but has recently been demonstrated in laboratories, are called negative index materials or left handed materials. In these materials, the energy propagation direct is opposite to the phase propagation direction.

Here we will consider the more general case of oblique incidence of an electromagnetic wave onto an interface. As shown in the following figure, a plane electromagnetic wave propagates along direction \mathbf{k}_i (wave vector direction) and meets an interface with norm $\hat{\mathbf{n}}$. The reflected wave and refracted wave propagates along the \mathbf{k}_r and \mathbf{k}_t directions, respectively. We call the plane formed by \mathbf{k}_i and $\hat{\mathbf{n}}$ as the plane of incidence, and the angle formed between $\hat{\mathbf{n}}$ and \mathbf{k}_i as the angle of incidence.



When an electric field is parallel to the plane of incidence, its conjugated magnetic field component, in this case pointing out of the paper, is perpendicular to the plane of incidence and is thus always parallel to the interface (refer to the figure).

In the plane wave expression obtained earlier in this lecture, we have

$$\mathbf{k} \cdot \mathbf{r} = (k \sin \theta_i, 0, k \cos \theta_i) \cdot (x, y, z) = k(x \sin \theta_i + z \cos \theta_i),$$

$$k = \frac{N\omega}{c_0} = N \frac{2\pi}{\lambda_0}.$$

Thus the incident, reflected, and transmitted electric fields can be expressed as,

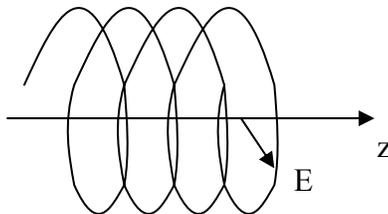
$$E_{//i} \exp \left[-i \left(\omega t - n_1 2\pi \frac{x \sin \theta_i + z \cos \theta_i}{\lambda_0} \right) \right],$$

$$E_{//r} \exp \left[-i \left(\omega t - n_1 2\pi \frac{x \sin \theta_r - z \cos \theta_r}{\lambda_0} \right) \right],$$

$$E_{//t} \exp \left[-i \left(\omega t - n_2 2\pi \frac{x \sin \theta_t + z \cos \theta_t}{\lambda_0} \right) \right],$$

where in the second equation the negative sign before $z \cos \theta_r$ indicates different propagation direction (upward in the z direction) from the incoming and refraction waves. The subscript “//” means that the electric field is polarized parallel to the plane of incidence for the sketched TM transverse wave.

Note: For EM waves, two transverse vibrating directions exist. In general, the resultant electric field vector will moves around in an ellipse instead of pointing in the same direction all the time. The path of the E vector is like a spiral in this situation.



Assuming there is no net surface charge or current on the interface ($z=0$), we can apply the continuity boundary condition to the vertical and tangential electric fields, respectively. In the x -direction, the tangential fields give

$$\begin{aligned} & \cos \theta_i E_{//i} \exp \left[-i\omega \left(t - \frac{n_1 x \sin \theta_i}{c_0} \right) \right] + \cos \theta_r E_{//r} \exp \left[-i\omega \left(t - \frac{n_1 x \sin \theta_r}{c_0} \right) \right] \\ &= \cos \theta_t E_{//t} \exp \left[-i\omega \left(t - \frac{n_2 x \sin \theta_t}{c_0} \right) \right] \end{aligned}$$

where the above equation is valid only when the exponents are equal because x can take any value. Thus we have

$$n_1 \sin \theta_i = n_1 \sin \theta_r = n_2 \sin \theta_t$$

which leads to the **Snell law** for reflection and refraction

$$\theta_i = \theta_r \quad \text{and} \quad n_1 \sin \theta_i = n_2 \sin \theta_t,$$

which yield

$$\cos \theta_i E_{//i} + \cos \theta_r E_{//r} = \cos \theta_t E_{//t}.$$

The magnitude of the magnetic field, which is pointing out of the paper, is related to the electric field by

$$H_y = \frac{n}{\mu c_0} E_{//}.$$

We can write the continuity of the tangential component of the magnetic field as

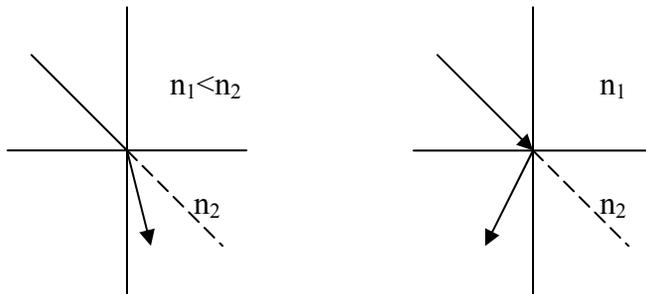
$$n_1 E_{//i} - n_1 E_{//r} = n_2 E_{//t}.$$

Based on the equations for electric and magnetic fields, we obtain the reflection coefficient, $r_{//}$, and transmission coefficient, $t_{//}$, for a TM wave as

$$r_{//} = \frac{E_{//r}}{E_{//i}} = \frac{-n_2 \cos \theta_i + n_1 \cos \theta_t}{n_2 \cos \theta_i + n_1 \cos \theta_t} = \frac{\sin 2\theta_t - \sin 2\theta_i}{\sin 2\theta_t + \sin 2\theta_i},$$

$$t_{//} = \frac{E_{//t}}{E_{//i}} = \frac{2n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_t}.$$

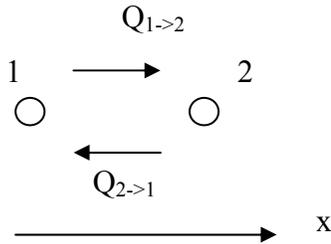
Note: (1) For an incident light shown as below (e.g. from air to water), in this case the refraction light will be bended and the object in the second media will looks higher if the observer is in the first media. (2) If $n_2 < 0$, $n_1 > 0$, the refraction light will be bended as the right figure. However, in nature no negative- n materials exist.



2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 13

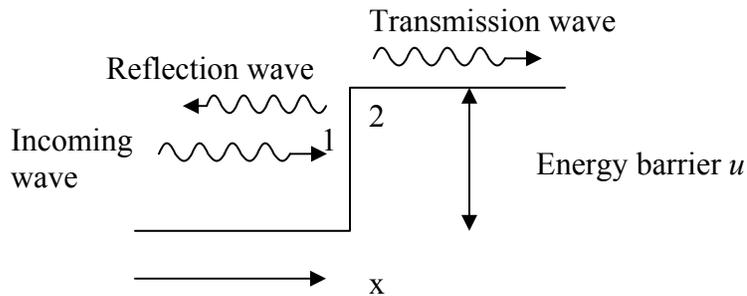
Review of previous lectures

1. Energy transport between two points

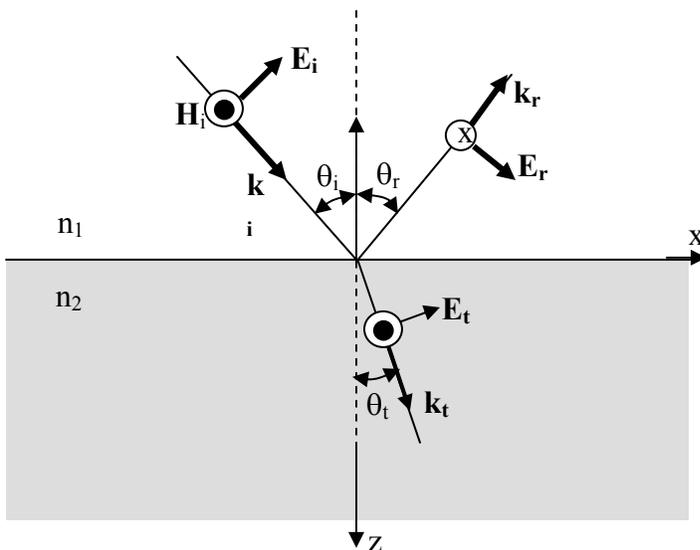


2. Plane waves & their interface reflection

We are interested in the wave energy at points 1 and 2 on two sides of the interface.



3. Oblique incidence of an electromagnetic wave onto an interface



Assuming flat interface, we have $\theta_i = \theta_r$. In last lecture, we have derived

$$r_{//} = \frac{E_{//r}}{E_{//i}} = \frac{-n_2 \cos \theta_i + n_1 \cos \theta_t}{n_2 \cos \theta_i + n_1 \cos \theta_t}, \quad t_{//} = \frac{E_{//t}}{E_{//i}} = \frac{2n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_t},$$

which are known as the **Fresnel coefficients** of reflection and transmission. For normal incidence ($\theta_i = 0$), similarity exists between case 2 and case 3 (see the following table).

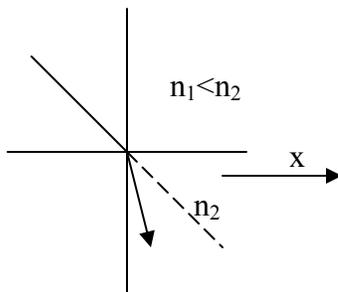
Electron propagation across a barrier	Normal incidence onto an interface
$r = -\frac{\Psi_r}{\Psi_i} = \frac{k_1 - k_2}{k_1 + k_2}$	$r_{//} = \frac{E_{//r}}{E_{//i}} = \frac{-n_2 + n_1}{n_2 + n_1}$
$t = \frac{2k_1}{k_1 + k_2}$	$t_{//} = \frac{2n_1}{n_2 + n_1}$
$R = -J_r / J_i = \left \frac{B}{A} \right ^2 = \left \frac{k_1 - k_2}{k_1 + k_2} \right ^2$	$R_{//} = \frac{S_{r,z}}{S_{i,z}} = \frac{S_r}{S_i} = r_{//} ^2$
$T = \frac{J_t}{J_i} = \frac{\text{Re}(k_2^*)}{\text{Re}(k_1^*)} t ^2$	$T_{//} = \frac{S_{t,z}}{S_{i,z}} = \text{Re} \left(\frac{n_2 \cos \theta_t}{n_1 \cos \theta_i} \right) t_{//} ^2$

Discussions

1) Critical & total internal reflection

A) $n_1 < n_2$

The Snell law is applied, $n_1 \sin \theta_i = n_2 \sin \theta_t$.



Note: The Snell law indicates the momentum conservation, or wavevector $k_{x1} = k_{x2}$ on the interface.

B) $n_1 > n_2$

Because the maximum angle of the refracted wave is $\theta_t = 90^\circ$, there exists an angle of incidence above which no real solution for θ_t exists. This **critical angle** happens when, according to the Snell law,

$$n_1 \sin \theta_c = n_2 \sin 90^\circ \quad \text{or} \quad \sin \theta_c = \frac{n_2}{n_1}$$

Above this angle, the reflectivity equals one, i.e., all the incident energy is reflected ($T=0$, $R=1$).

For an electromagnetic wave incident above the critical angle, the Snell law gives,

$$\sin \theta_t = \frac{n_1 \sin \theta_i}{n_2} > 1,$$

and thus,

$$\cos \theta_t = \sqrt{1 - \sin^2 \theta_t} = i \sqrt{\left(\frac{n_1 \sin \theta_i}{n_2}\right)^2 - 1} = ai.$$

In the wave function of the transmitted wave, the imaginary $\cos \theta_t$ leads to an exponential decay wave

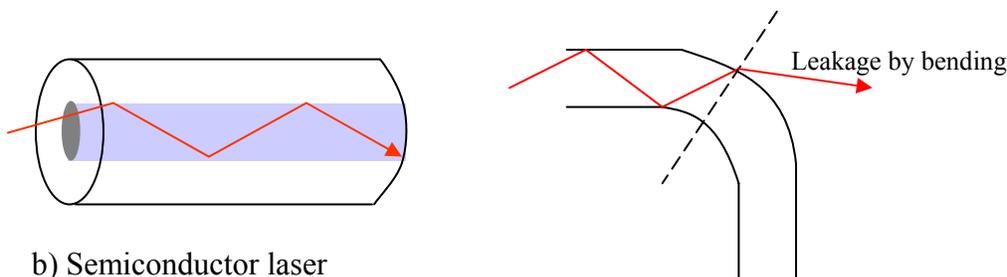
$$\begin{aligned} E_{//t} \exp \left[-i\omega \left(t - \frac{n_2 x \sin \theta_t + n_2 z \cos \theta_t}{c_o} \right) \right] \\ = E_{//t} \exp \left[-i\omega \left(t - \frac{n_2 x \sin \theta_t}{c_o} \right) - \frac{n_2 z \omega a}{c_o} \right], \end{aligned}$$

which is similar to the encountered evanescent wave.

Two applications:

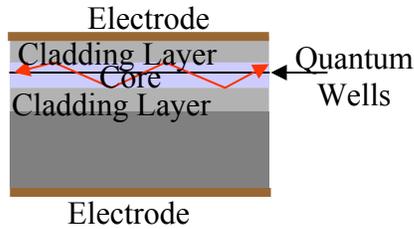
a) Optical fiber

An optical fiber has a core region and a cladding layer. The refractive index in the core region is higher than in the cladding layer. If light is launched into the fiber at an angle larger than the critical angle, the light will be bounced inside the core only without leakage, thus traveling a long distance along the fiber if the absorption coefficient of the core is small. However, the light can still escape the fiber core if we bend it.



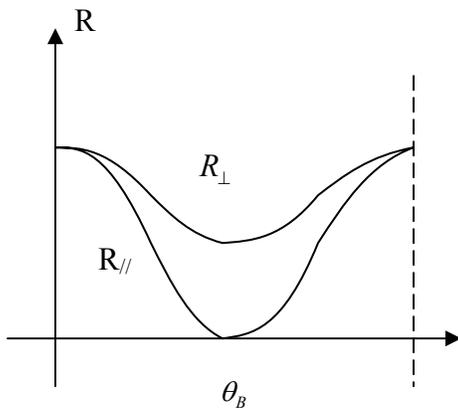
b) Semiconductor laser

In a semiconductor laser, light is emitted through electron-hole recombination inside the quantum well. The emitted light spreads over the core region and is confined by cladding layers that have a low refractive index than the core.



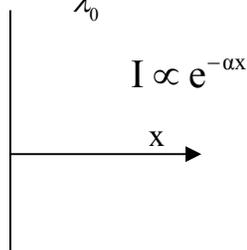
2) Brewster angle

When $\theta_i + \theta_t = \frac{\pi}{2}$, we have the electric wave $r_{//} = 0$, but the magnetic wave $r_{\perp} \neq 0$. This incident angle θ_B is called the **Brewster angle**. The corresponding reflectivity is drawn in the following figure.



3) Complex refractive index

For a complex refractive index $N = n + i\kappa$, the intensity of the wave decay as $I \propto e^{-\alpha x}$, where $\alpha = \frac{4\pi\kappa}{\lambda_0}$.



For real n_1 and complex N_2 , from $n_1 \sin \theta_i = N_2 \sin \theta_t$, we obtain complex θ_t and

$$\sin \theta_t = \frac{n_2 \sin \theta_i}{n_1} = a + bi, \quad \cos \theta_t = \sqrt{1 - \sin^2 \theta_t} = c + di.$$

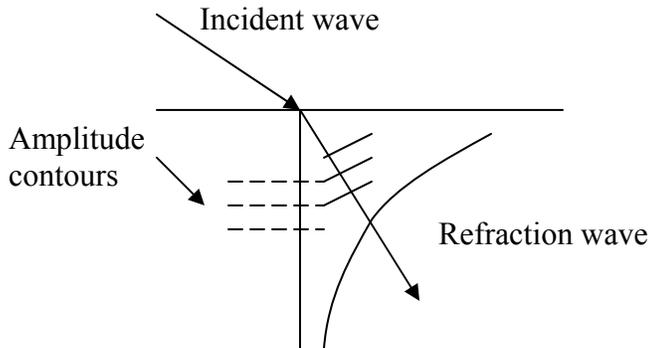
Thus the transmission wave is

$$E_{//t} \exp \left[-i\omega \left(t - \frac{n_2 x \sin \theta_t + (n_2 + i\kappa_2)(c + di)z}{c_0} \right) \right].$$

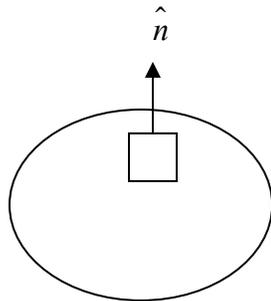
Let $c_1 + d_1 i = (n_2 + i\kappa_2)(c + di)$. Similar to the evanescent wave, the imaginary component

of the product leads to a decaying amplitude, while the real part contributes to the phase factor. The energy flow is still

$$\langle S \rangle = \frac{1}{2} \text{Re}(\bar{E} \times \bar{H}^*).$$



4) Acoustic waves



Recall the Newton's law

$$\bar{F} = m\bar{a},$$

where the force is $\bar{F} = \hat{n} \cdot \bar{\sigma}$, acceleration is $\bar{a} = \frac{d\bar{v}}{dt}$. Denote u as the displacement, i.e.,

$\bar{u} = \bar{x} - \bar{x}_0$. We have

$$\bar{v} = \frac{d\bar{x}}{dt} = \frac{d\bar{u}}{dt}.$$

In mechanics, we define the strain as

$$S_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right),$$

and the stress is

$$\bar{\sigma} = \bar{c} \cdot \bar{S}.$$

Note: \bar{c} is a fourth-order tensor and has 81 components.

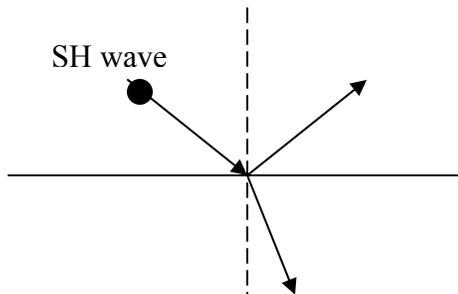
For phonons, we have two transverse waves and one longitudinal wave. The Poynting vector for acoustic waves is

$$\langle \bar{p} \rangle = -\frac{1}{2} \operatorname{Re}(\bar{v}^* \cdot \bar{\sigma}).$$

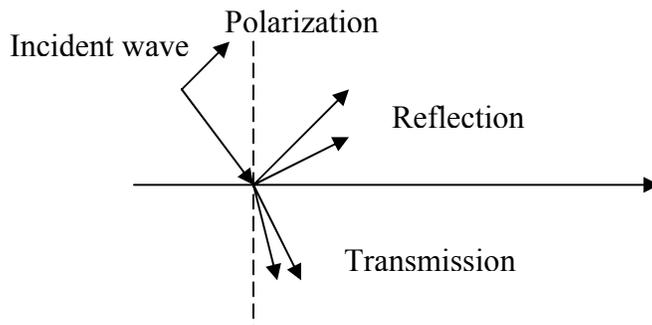
When the media is isotropic and the incident wave is a transverse wave with displacement polarized in the direction perpendicular to the plane of incidence (called a shear wave or SH wave), only one SH reflected and one SH transmitted wave are excited. At normal incidence, the acoustic reflectivity for a SH wave is

$$R_s = \left| \frac{Z_1 - Z_2}{Z_1 + Z_2} \right|^2,$$

in which the acoustic impedance Z (similar to refractive index in optics) is defined as $Z = \rho v$.



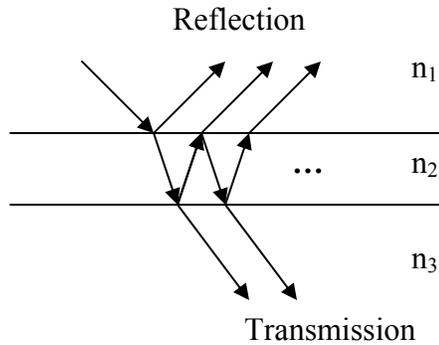
When the incident wave is polarized in the direction parallel to the plane of incidence, one longitudinal wave and one transverse wave are excited for both reflection and transmission waves. If the materials are anisotropic, one more transverse is excited for both reflection and transmission waves.



Note: The thermal resistance on the interface is very important for nanomaterials. Acoustic waves (or heat propagation) can be cut off by the interface, just as using a foil to cut off the radiation between two surfaces.

5.3 Wave propagation in thin films

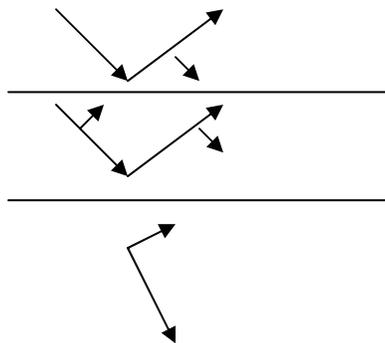
First let us consider the following structure. Many reflections and transmissions exist in this case. Summation is required to calculate the total reflection and transmission.



To avoid the summation of infinite series, two other methods are utilized:

1) Resultant wave method

In this method, the multiple reflection or transmission waves are combined into one in every material. In the following figure, we have four unknowns and four interfacial boundary conditions. The reflectivity and transmissivity can be determined on each interface.

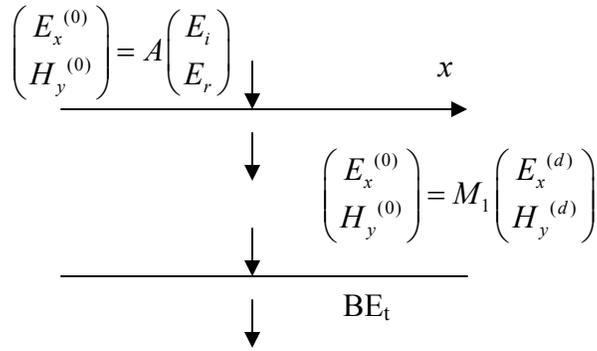


2) Transfer matrix method

The transfer matrix method combines all the waves (both forward and backward ones) in each medium into one wave, and uses a matrix to relate the electric and the magnetic fields between two different points inside a medium. Because the tangential components of the electric and the magnetic fields are continuous across the interface when there is no interface charge and interface current, the transfer matrix method can be easily extended to multilayers.

In the following figure, the x-component of the electric field $E_x^{(z)}$ and y-component of the magnetic field $H_y^{(z)}$ on the interface are related by 2×2 matrix A, 2×2 matrix M_1 , 2×1 matrix B. And we obtain

$$\begin{pmatrix} E_i \\ E_r \end{pmatrix} = A^{-1} M_1 B E_t .$$



For multilayers, M is determined by the chain rule as
 $M = M_1 M_2 \cdots M_n$,
 where n is the number of layers.

For a single layer of film, $\begin{pmatrix} E_i \\ E_r \end{pmatrix} = A^{-1} M_1 B E_t$ yields

$$r = \frac{E_r}{E_i} = \frac{r_{12} + r_{23} \exp[2i\varphi_2]}{1 + r_{12} r_{23} \exp[2i\varphi_2]},$$

$$\varphi = \frac{2\pi n_2 d \cos \theta_2}{\lambda_0}.$$

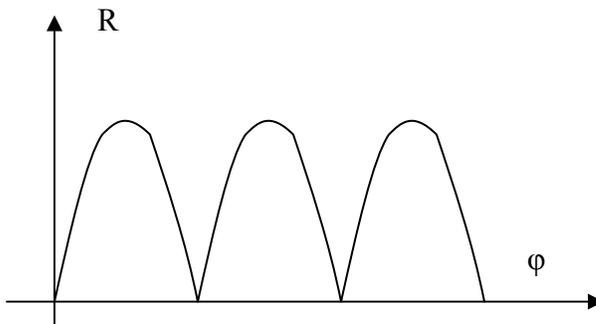
Thus for a nonabsorbing film, we obtain

$$R = |r|^2 = \frac{r_{12}^2 + r_{23}^2 + 2r_{12}r_{23} \cos 2\varphi_2}{1 + 2r_{12}r_{23} \cos 2\varphi_2 + r_{12}^2 r_{23}^2},$$

in which the cosine function indicates periodicity of R. This is just the interference effect.

Discussions:

(1) Periodic variation in R



Note: In microfabrication, the color of a thin film will change periodically according to the thickness, which is used to estimate the film thickness by eyes.

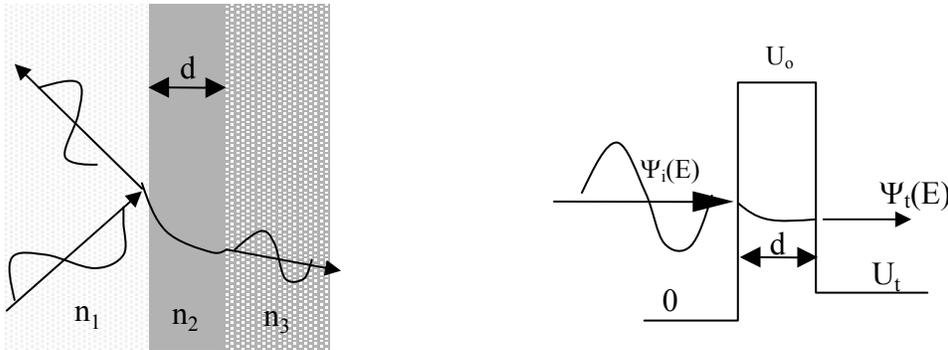
(2) Tunneling

Back to the case in which $\theta_i > \theta_{cr}$. We have

$$\cos \theta_t = \sqrt{1 - \sin^2 \theta_t} = i \sqrt{\left(\frac{n_1 \sin \theta_i}{n_2}\right)^2 - 1} = ai,$$

which causes decay in the second medium. If the second medium is very thin, a third medium attached to the second medium will have tunneling effects. The wave will be partly transmitted into the third medium before decaying off and the reflectivity $R \neq 1$ in this case.

Note: From $n_1 \sin \theta_i = n_2 \sin \theta_2 = n_3 \sin \theta_3$, in this case θ_1, θ_3 are real numbers, while θ_2 is imaginary number.



Similar phenomena happen to the electron propagation across a barrier. In the right figure, we can see tunneling happens when the barrier is thin. The transmissivity is

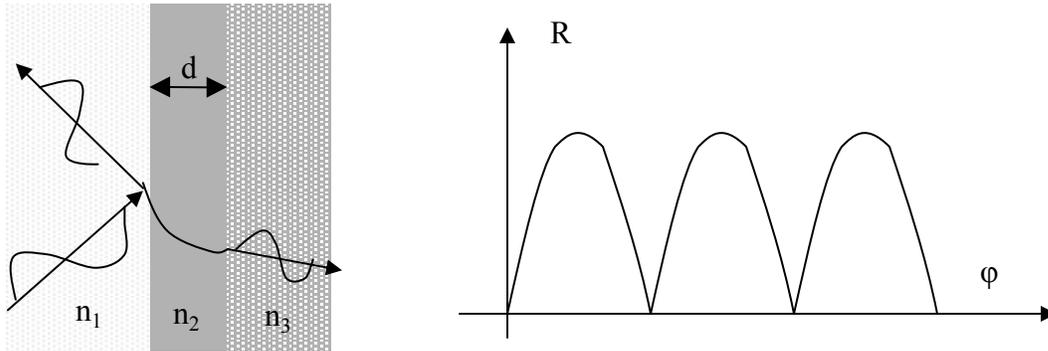
$$\tau = \frac{4E(U_0 - E)}{4E(U_0 - E) + U_0^2 \sinh^2 \left[\sqrt{2m(U_0 - E)}d / \hbar \right]}$$

or

$$\tau \approx \frac{16E(U_0 - E)}{U_0^2} \exp \left[-2\sqrt{2m(U_0 - E)}d / \hbar \right] = \frac{16E(U_0 - E)}{U_0^2} e^{-2|k_2|d}.$$

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 14

Review of last lecture

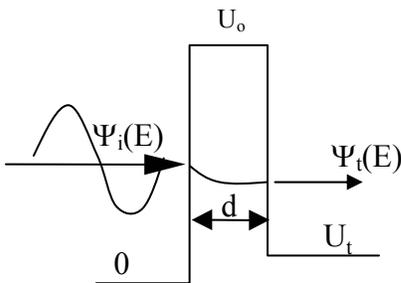


In lecture 13, we talked about tunneling through a thin film, which can also happen to general EM waves and acoustic waves. If the second medium is very thin, a third medium attached to the second medium will have tunneling effects. The wave will be partly transmitted into the third medium before decaying off and the reflectivity $R \neq 1$ in this case.

Define $\phi = 2\pi nd \cos \theta / \lambda_0$. For a thin film, the reflectivity is a periodic function of ϕ , as shown in the right figure. In the widely used coating technique, the color of a thin film will change periodically according to the thickness, which is used to estimate the film thickness by naked eyes.

We also talked about the tunneling through an energy barrier presented in the following figure. The barrier region can be vacuum in this case. The transmissivity is approximated as

$$\tau \approx \frac{16E(U_0 - E)}{U_0^2} \exp[-2\sqrt{2m(U_0 - E)}d/\hbar] = \frac{16E(U_0 - E)}{U_0^2} e^{-2|k_2|d}$$



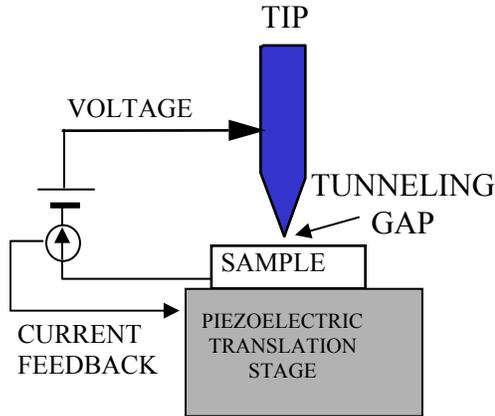
Now let us estimate the d value for a 1eV energy barrier. The k_2 value is

$$k = \sqrt{\frac{2m(u - E)}{\hbar^2}} \sim \sqrt{\frac{2 \times 9.1e-31 \times 1.6e-19}{1e-68}} \sim 4e9 \text{ m}^{-1},$$

and the characteristic length for d is $1/k \sim 1 \text{ \AA}$.

Applications of tunneling

1) Scanning tunneling electron microscope (STM)

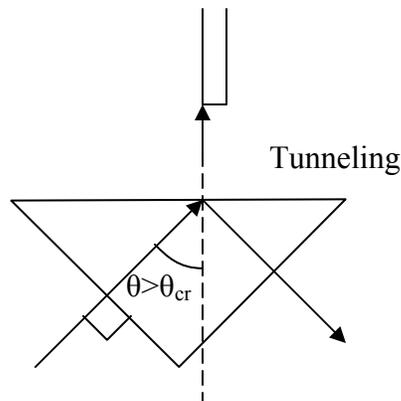


The **tunneling phenomena** are the basis of several inventions that led to several Nobel prizes including the tunneling diode by Esaki (1958) and the scanning tunneling electron microscope (STM) (Binnig and Rohrer, 1982). As shown in the above figure, in a STM a sharp tip is brought in close proximity with a conducting surface but not contacting the surface. The piezoelectric stage can adjust the distance between the tip and sample with subatomic accuracy. Under an applied voltage, electrons tunnel through the vacuum gap and create a current in the loop. The current is extremely sensitive (sub-angstrom) to the separation between the tip and the contact because k_2 is on the order of $\sim 1 \text{ \AA}^{-1}$ and the transmissivity changes exponentially according to d . As the tip is scanned over the sample, different region has different potential barrier or different heights. By using the current as a feedback signal, one can map the electronic wavefunction surrounding individual atoms or the surface roughness.

2) Other microscopes

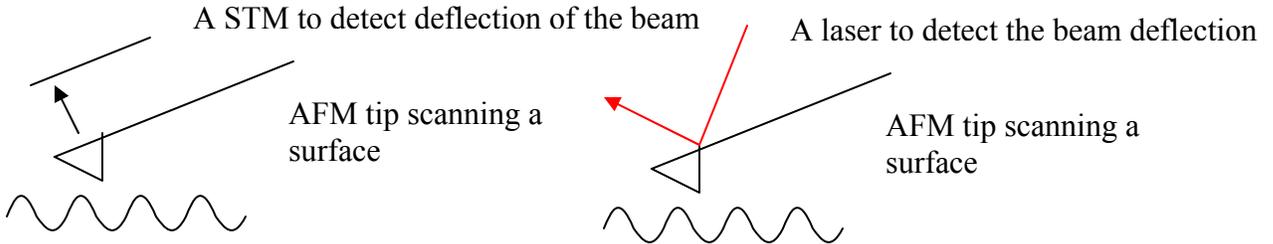
A) Photon scanning tunneling microscope

In such a microscope, the incident angle on the scanned surface is larger than the critical angle and thus it will be totally reflected in normal cases. However, when a scanning probe approaches the other side of the surface, there will be tunneling between the surface and the probe and the topography is obtained as in a STM.



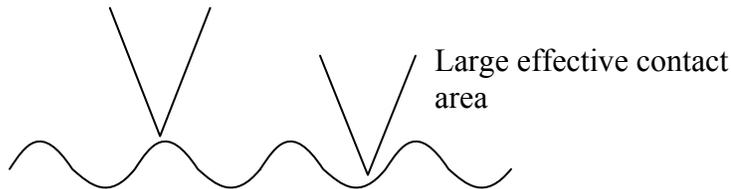
B) Atomic force microscope (AFM)

A STM cannot be used to scan a dielectric surface because the surface needs to be conductive to provide tunneling electrons. To deal with a nonconductive surface, an AFM is invented in Stanford University.



The origin idea of an AFM is shown in the left figure. A diamond is attached to an Al-film to form a scanning probe. A STM is mounted on top of the Al film to detect its deflection. Since the spring constant between atoms is much larger than that of the beam (Al film), in scanning the beam will bend according to the surface topography but the atoms will not be affected. In this way the surface image is obtained. In the current AFM, the STM is replaced by a laser beam, as shown in the right figure.

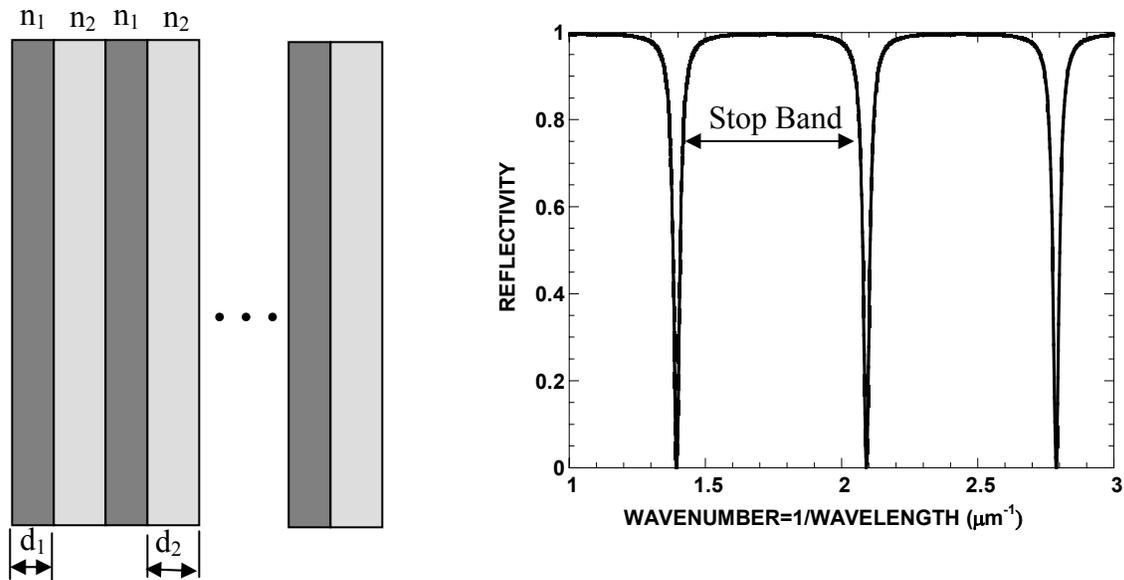
By mounting a thermocouple onto the top of the tip, a scanning thermal microscope (SThM) is created. In thermal imaging, the probe must contact the surface all the time. It should be noted that the thermal imaging are not as precise as topography imaging. In the following figure, the effective contact area between the tip and surface is not constant and will affect the image quality. Additionally, thermal imaging normally requires high vacuum to avoid air conduction and convection effects.



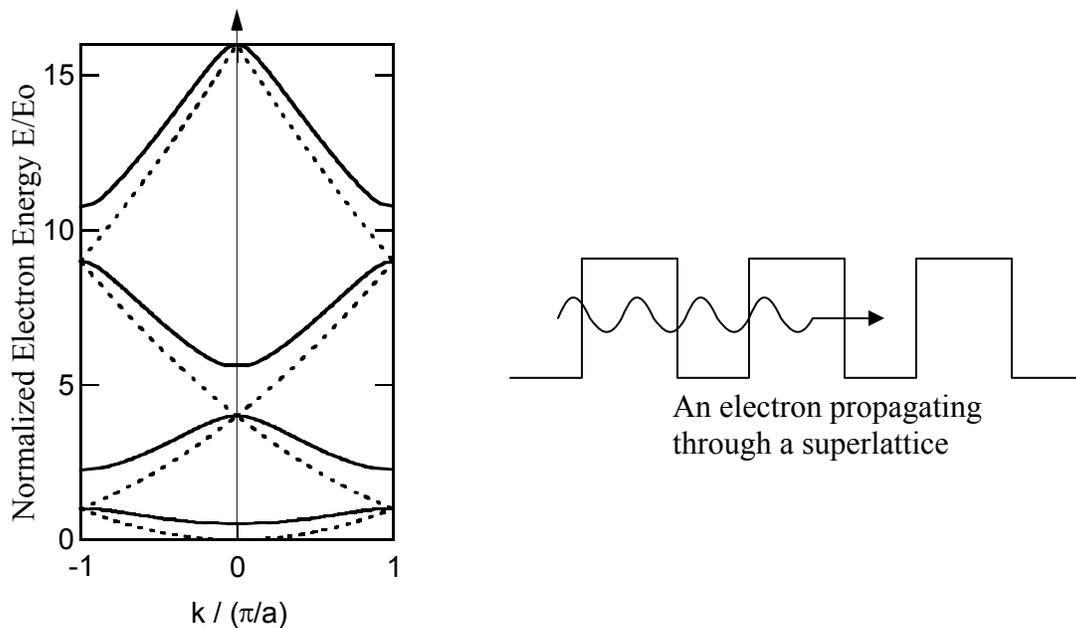
5.4 Bragg reflector

In practice, the reflectivity and transmissivity of multilayers can be controlled quite accurately with various thin-film deposition techniques and the possibility in controlling spectral and directional properties is large. In the following left figure, we present a **Bragg reflector** that is made from two alternating layers of thin films. Each layer has a thickness equal to the one-quarter of the light wavelength inside the film. Although at one interface, the reflectivity between the two materials may be small, the coherent superposition of the reflected fields can create a reflectivity that is close to 100%. Such Bragg reflectors are used in coatings for highly reflectivity mirrors at specific required wavelength, such as for lasers and X-rays. The right figure gives an example of the reflectivity of a quarter wavelength mirror, similar to these used in a special semiconductor laser structures called the vertical-cavity surface-emitting lasers (Koyama et al., 1989; Walker et al., 1993). The reflectivity in certain spectrum regions can reach

100% reflectivity, meaning no electromagnetic fields of corresponding wavelength exist inside the reflector. These spectral regions, called stop bands, occur when the round-trip phase difference through one period (two layers) equal $2 \ell \pi$.

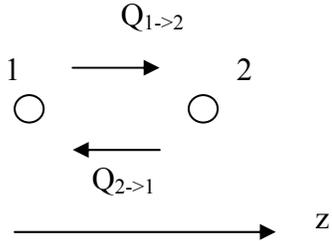


It is interesting to compare the stop band with electron band gap in semiconductors. In the latter case, no electrons exist in the forbidden energy levels, while in the reflector no photons exist in the stop band. The idea of forbidden energy in semiconductors is also utilized to filter electrons in superlattice, whose inventor, Esaki won a Nobel Prize for his work on double-barrier tunneling effect.



Note: The research in 3D semiconductors triggered the work on 1D superlattice, and now the research go back to 3D cases in 3D Bragg reflector, also called photonic crystals. Photonic band is formed within such a crystal and filter out radiation with unwanted wavelengths.

5.5 Landauer formalism



Now let us go back to the energy exchange between two points, which is expressed as

$$q_{1 \rightarrow 2} = \sum_{p=1}^{3m} \left[\frac{1}{V_1} \sum_{k_{x1}=-\infty}^{\infty} \sum_{k_{y1}=-\infty}^{\infty} \sum_{k_{z1}=0}^{\infty} v_{z1} E_1 \tau_{12} f(E_1, T_{e1}) \right],$$

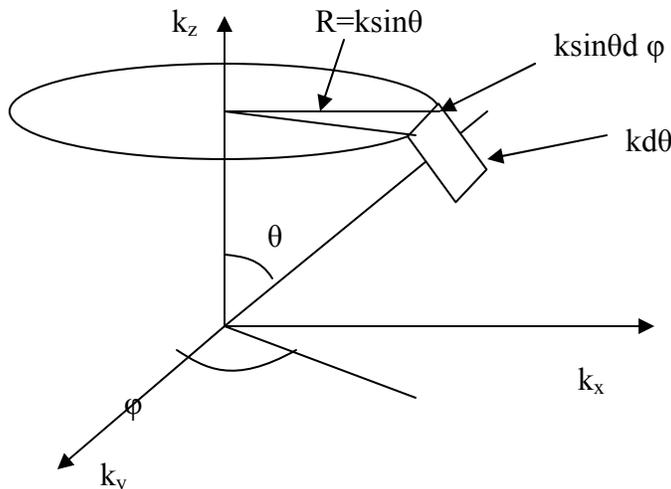
where T_{e1} represents the temperature of the phonons coming towards the interface and $f(E_1, T_{e1})$ is the Bose-Einstein distribution for phonons at T_{e1} , and τ_{12} is the phonon transmissivity from medium 1 into medium 2. The unit of energy flux q is W/m^2 , so we divide $v_{z1} E_1$ (J m/s) by the volume on the right side.

Similar to solving μ in lecture 11, here we can change the summation into integral form.

$$q_{1 \rightarrow 2} = \sum_{p=1}^{3m} \left[\int_V v_{z1} E_1 \tau_{12} f(E_1, T_{e1}) \left(\frac{dk_{x1} dk_{y1} dk_{z1}}{(2\pi/L)^3} \frac{1}{V_1} \right) \right],$$

in which

$$\frac{dk_{x1} dk_{y1} dk_{z1}}{(2\pi/L)^3} \frac{1}{V_1} = \frac{4\pi k^2 dk}{8\pi^3} = \frac{4\pi k^2}{8\pi^3} \frac{dk}{dE} dE = D(E_1) dE.$$



However, we also notice that velocity v_{z1} depends on k_{x1} and transmissivity τ_{12} depend on (k_{x1}, k_{y1}, k_{z1}) . To deal with this, first we note

$$dk_{x1} dk_{y1} dk_{z1} = k^2 \sin \theta dk d\theta d\varphi = k^2 dk d\Omega = k^2 \frac{dk}{dE} dE d\Omega,$$

where the solid angle $d\Omega = \frac{k^2 \sin \theta d\theta d\varphi}{k^2} = \sin \theta d\theta d\varphi$ in the spherical coordinates. Thus

$$q_{1 \rightarrow 2} = \sum_{p=1}^{3m} \left[\int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin \theta d\theta v_{z1} E_1 \tau_{12} f(E_1, T_{e1}) \frac{k^2}{8\pi^3} \frac{dk}{dE} dE \right].$$

For isotropic materials, $\frac{k^2}{8\pi^3} \frac{dk}{dE} = \frac{D(E)}{4\pi} = D'(E)$ because the solid angle of whole sphere is 4π (or $\frac{D(E)}{4\pi} d\Omega dE = D(E) dE$).

The total energy flux is defined as

$$q_{12} = q_{1 \rightarrow 2} - q_{2 \rightarrow 1} = q_{1 \rightarrow 2}(T_{e1}) - q_{2 \rightarrow 1}(T_{e2}) = q_{1 \rightarrow 2}(T_{e1}) - q_{1 \rightarrow 2}(T_{e2}),$$

where we use $q_{1 \rightarrow 2}(T_{e2}) = q_{2 \rightarrow 1}(T_{e2})$ in the last step. This is obtained by considering the equilibrium status $q_{12} = 0$ and $T_{e2} = T_{e1}$. In nonequilibrium situations, this relationship still holds true because we have the same expression for $q_{2 \rightarrow 1}(T_{e2})$. This idea is comparable to the Kirchoff's Law in radiation, in which we consider the energy exchange with a blackbody surface to derive $\varepsilon(\lambda, \theta) = \alpha(\lambda, \theta)$.

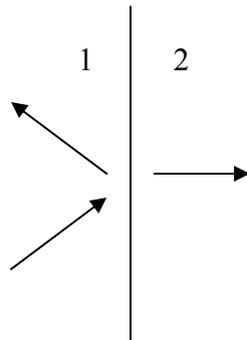
From above equations, we finally obtain the Landauer formalism

$$q_{1 \rightarrow 2} = \sum_{p=1}^{3m} \left[\int_S d\Omega \int v_{z1} E_1 \tau_{12} [f(E_1, T_{e1}) - f(E_1, T_{e2})] \frac{D(E_1)}{4\pi} dE \right],$$

where the solid angle is integrated over the hemisphere.

Examples:

1) Interfacial thermal conductance



Consider the heat conduction through the interface. When $T_{e1} \approx T_{e2}$, Taylor expansion

$$f(E_1, T_{e1}) - f(E_1, T_{e2}) = \frac{\partial f}{\partial T} (T_{e1} - T_{e2})$$

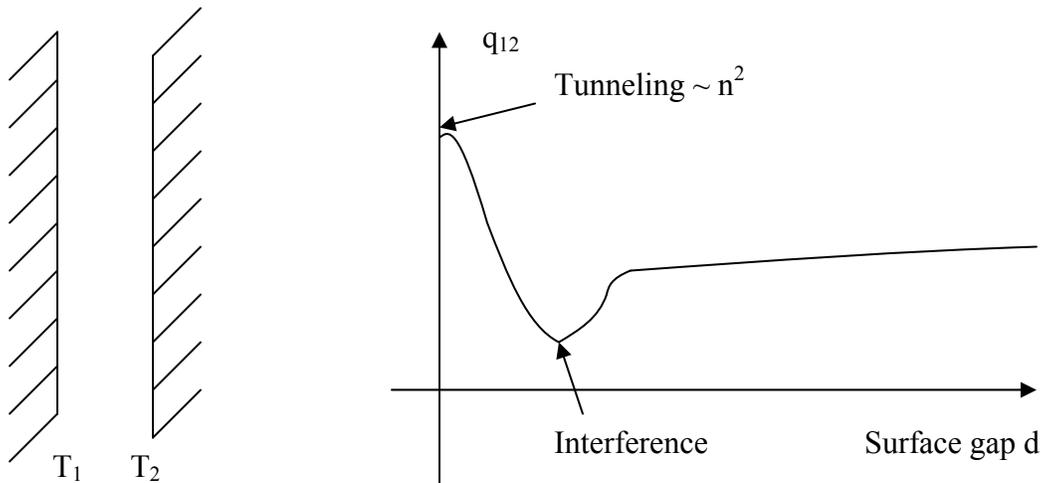
applies in the Landauer formalism and we obtain

$$q_{12} = K(T_{e1} - T_{e2}) = \frac{T_{e1} - T_{e2}}{R_{12}},$$

where K is the contact thermal conductance ($\text{W}/\text{m}^2 \text{K}$), R_{12} is the contact thermal resistance ($\text{m}^2 \text{K}/\text{W}$).

Note: (1) The value of τ_{12} ranges from 0 to 1. In the Landauer formalism, $E_1 \frac{\partial f}{\partial T} D(E) dE$ has the physical meaning of volumetric specific heat C . For normal materials, the magnitude of C is 10^6 (mass-based specific heat can differ a lot) and v is on the order of 10^3 . Thus $K \sim v z_1 C \sim 1e3 \times 1e6 = 1e9 \text{ W}/\text{m}^2 \text{K}$. (2) Based on one-dimensional heat conduction, $R = \frac{L}{kA}$ leads to the equivalent thickness of the interface is $0.1 \mu\text{m}$ if k is on the magnitude of 10^2 . (3) For electron flow, the chemical potential μ will drop across the interface instead of the temperature. If the contact resistance is large, electrons will be reflected back. (4) The interfacial thermal conductance is dominant for nanostructures such as superlattice. In applications, the contact resistance between a heat sink and CPU chip is a big problem for heat release.

2) Radiation tunneling



For two blackbody surfaces separated by a gap d , tunneling will occur when d is small. Beyond that region the energy flux follows the fourth-power law, $q_{12} = \sigma(T_1^4 - T_2^4)$. Due to tunneling, the maximum heat flux can exceed the blackbody radiation in vacuum. If the refractive index is n , we have $D(E) \sim n^3$ and $v \sim 1/n$ (Snell's law). Thus the heat flux is proportional to n^2 .

When a surface wave (decays exponentially on both sides of the surface) exists, the maximum heat flux can be higher than n^2 due to tunneling, because the density of states of the surface wave is much higher.

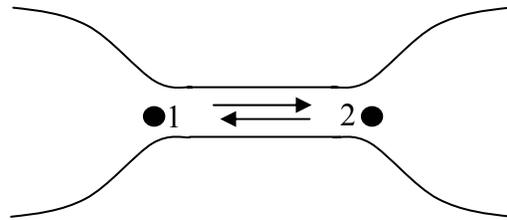
3) Quantum conductance (Universal conductance)

In the following figure, electrons flow between two points 1 and 2. Under the assumption of $\tau_{12}=1$ in the Landauer formalism, each quantum state corresponds to one “channel” of

energy carrier transport. For electrons, the conductance quantum is $K_e = \frac{2e^2}{h}$ (h is the

Planck’s constant), while for phonons $K_p = \frac{\pi^2 k_B T}{3h}$. It should be noted the conductance is

independent of materials. The conductance looks like stairs in this case.

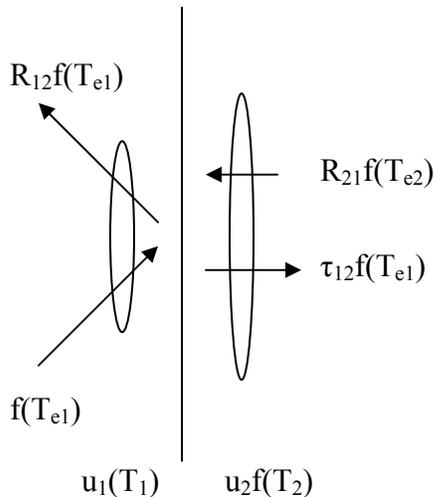


Discussion: What is temperature?

The Landauer formalism is not applicable to special cases where two media are identical and have perfect contact (i.e. merging into a whole bulk material). In these situations, the relationship $q_{12} = K(T_{e1} - T_{e2}) = \frac{T_{e1} - T_{e2}}{R_{12}}$ indicates that $T_{e1} \neq T_{e2}$ for any nonzero heat flux.

Obviously this violates the continuity of temperatures in a crystal. The paradox is explained by checking the temperature definition carefully. The local temperature T_{e1} is not equal to the temperature T_1 (measured beyond one mean free path) calculated from the local internal energy $u(T_1)$ in the small region around point 1.

Note: The relationship $q_{12} = K(T_1 - T_2)$ is consistent with the Fourier’s law.



2.57 Nano-to-Macro Transport Processes

Fall 2004 - Lecture 15

Guest lecture by Prof. Dresselhaus

1. Outline

- Overview
- Synthesis
- Structures and Symmetry
- Electronic Properties
- Transport Properties
- Phonon Properties
- Resonant Raman Effect
- Applications

2. Unique Structure and Properties of Single Wall Carbon Nanotubes (SWNTs)

A SWNT can be viewed as a cylinder formed by rolling up a graphene sheet.

Image removed for copyright reasons.

A SWNT is an ideal model of 1D systems for nanoscience. It has the following interesting characteristics:

(1) Size

SWNTs are nanostructures with dimensions of ~ 1 nm diameter (~ 20 atoms around the cylinder). The smallest SWNT has a diameter of only 0.4 nm.

(2) Electronic Properties

They can be either metallic or semiconducting depending on the diameter or orientation of the hexagons.

(3) Mechanical

SWNTs have very high strength, Young's modulus, and good properties on compression and extension.

SWCNTs can be used to make heat pipes and electromagnetic antennas. Its structure can be determined by single nanotube (as one molecule) spectroscopy. Due to the unique properties, currently many applications are being attempted worldwide for CNTs.

3. Synthesis

Three methods are utilized to grow CNs:

(1) Arc Discharge

In the following figure, two graphite rods (5-20 mm in diameter) are used as the cathode and anode, between which arcing occurs when 50-120A DC is supplied. By electron collision into the anodic rod, carbon clusters from the anodic graphite rod are condensed on the surface of the cathodic graphite rod and carbon nanotubes are formed along with other products.

Image removed for copyright reasons.

Y. Saito *et al*, *Phys. Rev.* 48 1907 (1993)

(2) Laser Ablation

The experimental setup is shown as following, where Nd-Yb-Al-garnet Laser is utilized at 1200 °C.

Image removed for copyright reasons.

A. Thess *et al*. *Science* 273 483 (1996)

(3) Chemical Vapor Deposition (CVD) method

For the laser ablation method, we always get many twisted “wires”, which are bundles of SWCNs and thus difficult to use. Isolated single wall carbon nanotubes can be grown by the Chemical Vapor Deposition (CVD) method. By depositing catalyst on the specified positions, we can control the location of grown CNs. This brings tremendous convenience to the research. Some work has also been conducted to control the average diameter and diameter distribution by changing the catalysts and furnace temperature (H. Kataura et al., 2000).

Image removed for copyright reasons.

N. Wang et al. Nature 408, 50 (2000)

The above figure shows the smallest SWNT with a diameter 0.42 nm. It is a (5,0) zigzag nanotube and has metallic electronic structure. Under 15 K it becomes a superconductor.

Image removed for copyright reasons.

Peapod Empty SWNT

H. Kataura et al, unpublished (2001)

A new material, fullerene-peapods can also be made based on SWCNs. Baking fullerenes with CNs in a quartz ampoule at 650 °C for 2 ~ 6 h, the fullerenes can fit into CNs and form this new structure. From the above figure, we can see its electronic structure is similar to empty SWNTs especially near the Fermi level. By Heating at higher temperatures, C₆₀ fullerenes inside the nanotubes can merge and form double wall carbon nanotubes (DWNTs).

4. Structures and Symmetry

Image removed for copyright reasons.

Chiral Vectors : (n,m)

R. Saito *et al.*, *Phys. Rev.* **B46**, 1804 (1992)

$$C_h = na_1 + ma_2 \equiv (n, m)$$

a_1, a_2 : primitive lattice vectors

$$T = t_1a_1 + t_2a_2 \equiv (t_1, t_2)$$

$$t_1 = \frac{(2m+n)}{d_R}, t_2 = -\frac{(2n+m)}{d_R}$$

$$d_R = \text{gcd}(2n+m, 2m+n)$$

$$d_t = \frac{L}{\pi} = \frac{a\sqrt{n^2 + nm + m^2}}{\pi}$$

$$L = |C_h|$$

Although there can be other ways to define the primitive lattice vectors in a hexagon sheet, the method presented in the above figure is widely used in all current publications. The chiral vector (equator of nanotube) is the vector OA or C_h in the figure; the translational vector of 1D material is vector OB and marked as T. Here C_h denotes the perimeter of the SWCN, while the translational vector is along the axis direction of the rolled up SWCNs. For one Unit Cell OAB'B, C_h and T are always expressed with the lattice vectors a₁ and a₂. Three special CN structures are defined as

- Symmorphic (mirror symmetry)
 - Armchair Nanotube (n,n), n=m
 - Zigzag Nanotube (n,0), m=0
- Non-Symmorphic (axial chirality)
 - Chiral Nanotube (n,m), n≠m

In the above figure, the chiral angle θ ($0 < \theta < \pi/6$) is defined as the included angle between C_h and a₁. The diameter d_t of SWNTs can be calculated from the length of C_h.

Image removed for copyright reasons.

P. Kim et al., PRL 82, (1999) 1225.

**J. W. G. Wildoer et al,
Nature, 391 (1998) 59**

Using a STM, the structure of a SWNT can be observed. In the right above figure, we can barely see the hexagons. SWNTs can be metallic or semiconductive. The transition happens when the conduction band and valence band touch each other at six K points in the k space. The energy contours are also drawn in the left above figure.

5. Electronic Properties

A unit cell and the corresponding Brillouin zone are drawn in the following two figures. The point K is at the corner of the hexagon, while M is the edge midpoint.

Image removed for copyright reasons.

P. R. Wallace, *Phys. Rev*, **71** 622 (1947).

Image removed for copyright reasons.

R. Saito *et al.*, *Phys. Rev.* **B46**, 1804 (1992)

In the above three cases, an energy gap appears in a (10,0) SWNT, indicating semiconductor. The density of states depends on the chirality of the SWNT. The general law to judge the electronic property (metallic, semiconductive) is summarized in the embedded equation of the following figure.

Image removed for copyright reasons.

$$n - m = \begin{cases} 3p & \text{metal} \\ 3p \pm 1 & \text{semiconductor} \end{cases}$$

R. Saito *et al.*, *Appl. Phys. Lett.* **60**, 2204 (1992)

The width of DOS split depends on the chirality of the SWNT. In the following figures, we find the Zigzag SWNT exhibits splitting DOS compared with an Armchair SWNT.

Image removed for copyright reasons.

(a) Metal

(b) Semiconductor

R. Saito *et al.*, *Phys. Rev.* **B61**, 2981(2000)

The wave vector k for one-dimensional carbon nanotubes is shown in the two-dimensional Brillouin zone of graphite hexagon. In the direction of K_1 , discrete k values are obtained by periodic boundary conditions for the circumferential direction of the carbon nanotubes, while in the direction of the K_2 vector, continuous k vectors are shown in the one-dimensional Brillouin zone. For metallic nanotubes, the bold line intersects a

K point (corner of the hexagon) at the Fermi energy of graphite. For the semiconductor nanotubes, the K point always appears one third of the distance between two bold lines and no DOS splitting occurs for any chirality.

6. Transport Properties

When two SWNTs with different atom numbers on the circumference meets, a pair of pentagon and heptagon will form at the junction and thus makes smooth transition between different geometries. Junctions such as a diode are measured in experiments.

Image removed for copyright reasons.

S. Iijima, NEC Symp.(1992) (a) Metal-metal (b) Metal-semiconductor

The following figures are the first electrical measurement of carbon nanotube, where the gate voltage is used to change the electronic structure of a SWNT.

Image removed for copyright reasons.

S.J. Tans *et al.* *Nature*, 393, 49 (1998)

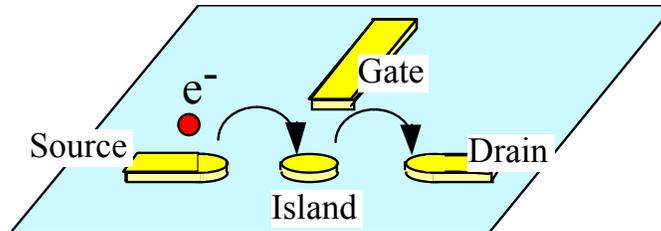
Resonant tunneling transport is demonstrated in the left bottom figure. The energy levels in a SWNT can be compared with a particle in a box. By changing the applied gate voltage, the energy level can be shifted and tunneling will happen at discrete levels. In addition, quantized conductance is also observed in this experiment, which is a multiple of

$$G_0 = \frac{2e^2}{h} = (12.9\text{k}\Omega)^{-1}$$

Image removed for copyright reasons.

W. Liang Harvard Univ.

Ballistic transport of electrons in a metallic carbon nanotube is also measured. For metallic SWNTs, the Fermi level just across the intersection point of two bands.



The above figure demonstrates the idea of CNT single electron transistor by AFM nicking. The nanotube is broken at two locations and electrons will transfer from source to drain under specified applied tip voltage.

7. Phonons and Raman Spectroscopy

The Raman Spectroscopy is the major characterization method for SWNTs. It is non-destructive, contactless measurement and can be conducted in air at ambient pressure. Additionally it is quick (1 min) and accurate in energy.

The Raman Spectroscopy measures the wavelength and intensity of inelastically scattered light from molecules. When light is scattered, a small fraction of light is scattered at optical frequencies different from, and usually lower than, the frequency of the incident photons. This inelastic scattering can occur with a change in vibrational, rotational or electronic energy of a molecule. Since Resonant Raman Effect is diameter selective (also chirality dependent), it is used to determine the size of SWNTs.

$$N = \frac{2(m^2 + n^2 + nm)}{d_R}$$

$$d_R = \begin{cases} 3d & \text{if } n - m = 3d \cdot p \\ d & \text{otherwise} \end{cases}$$

Image removed for copyright reasons.

“Physical Properties of Carbon Nanotubes” R. Saito, G. Dresselhaus, and M.S. Dresselhaus, Imperial College Press, (1998)

2N carbon atoms
6N phonon modes (three directions of vibrations)

Image removed for copyright reasons.

Phonon modes -- (10,10) Armchair
R.Saito *et al. Phys. Rev.* **B57** (1998) 4145

In the following figures, the result for SWNTs is presented.

Image removed for copyright reasons.

A.M. Rao et al, *Science* **275** (1997) 187

In the following figures, the PL absorption and emission spectra in S-SWNTs provides E_{11}^S , E_{22}^S , and additional information about the SWNT electronic structure. Compared to resonance Raman Spectroscopy, PL is a complementary technique to characterize a large number of semiconducting SWNTs in solution.

Image removed for copyright reasons.

M. O'Connell et. al. *Science* **297** 593 (2002)

8. Challenges for Carbon Nanotube Synthesis and Separation

It is important to control synthesis process to produce tubes with the same diameter and chirality (n,m). Before the control of synthesis process is achieved, we need to develop effective separation methods to separate metallic SWNTs from semiconducting ones by

diameter and chirality. For applications, it is significant to develop method for large-scale, cheap synthesis, and improve nanotube characterization and manipulation.

Separation processes reported till now include:

- (1) Precipitation of SWNTs non-covalently functionalized with ODA
- (2) Ion-exchange liquid chromatography of ssDNA wrapped SWNTs
- (3) Alternating current dielectrophoresis in an aqueous SWNT suspension
- (4) Selective functionalization with diazonium salts
- (5) Centrifugation after addition of diluted bromine

It should be noted that many of these processes separate SWNTs by diameter in addition to metallicity.

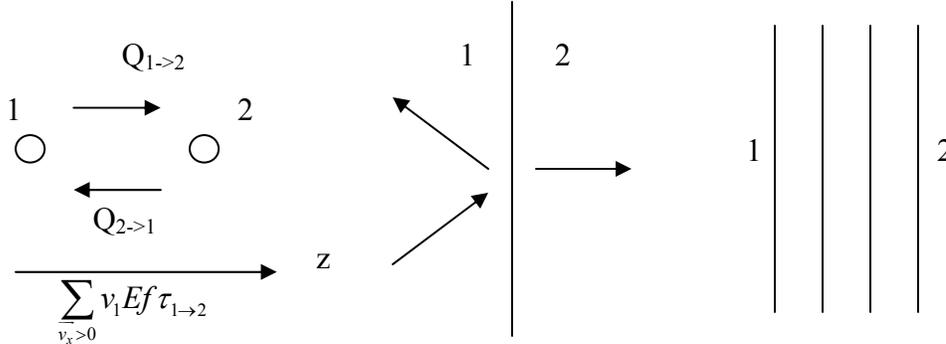
9. Applications

The wide applications of SWNTs include:

- (1) STM/AFM tips (advanteous in scanning sharp-trench topography)
- (2) Direct Analysis of DNA
- (3) Semiconductor devices
- (4) Field Emitters
- (5) Filler for enhancing lifetime in Li ion batteries (to eliminate the electrode gap change in the charging and recharging processes)
- (6) Filler for enhancing conductivity of polymer composites
- (7) Ultimate strong fiber (space elevator made of SWNTs)
- (8) Hydrogen Storage for fuel cells (as in the following figure)

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 16

Review of last lecture

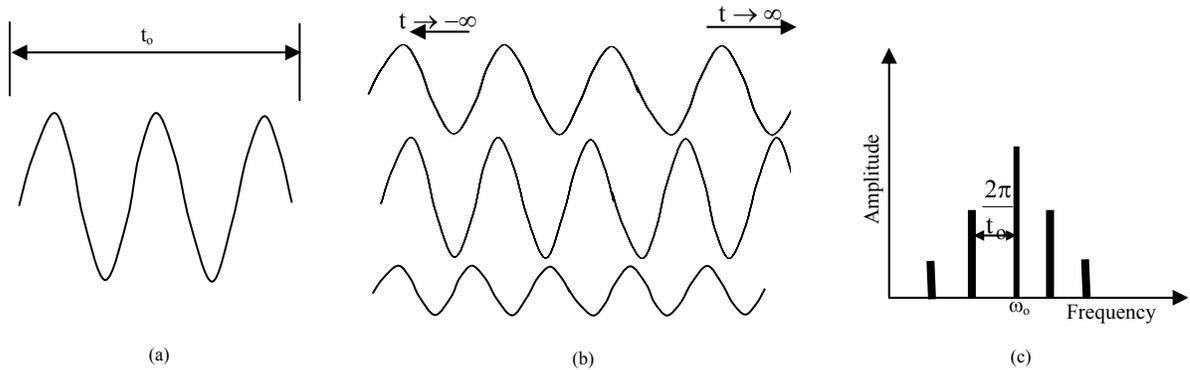


In Lecture 14, we have discussed the energy exchange between two points. We have spent time on calculating the transmissivity $\tau_{1 \rightarrow 2}$ for cases in the right two figures. The velocity v_1 is the group velocity, not the phase velocity determined by $v = \omega / k_x$ (derived from constant phase $\Phi = \omega t - k_x x$.)

Consider a plane traveling along x-direction

$$\mathbf{A}e^{-i(\omega t - kx)}$$

The velocity $v = \omega / k_x$ is NOT the speed of signal or energy propagation. We see that the plane wave represented by above equation extends from minus infinite to plus infinite in both time and space. There is no start or finish and it does not represent any meaningful signal. In practice, a signal has a starting point and an ending point in time. Let's suppose that a harmonic signal at frequency ω_0 is generated during a time period $[0, t_0]$, as shown in the following figure (a).



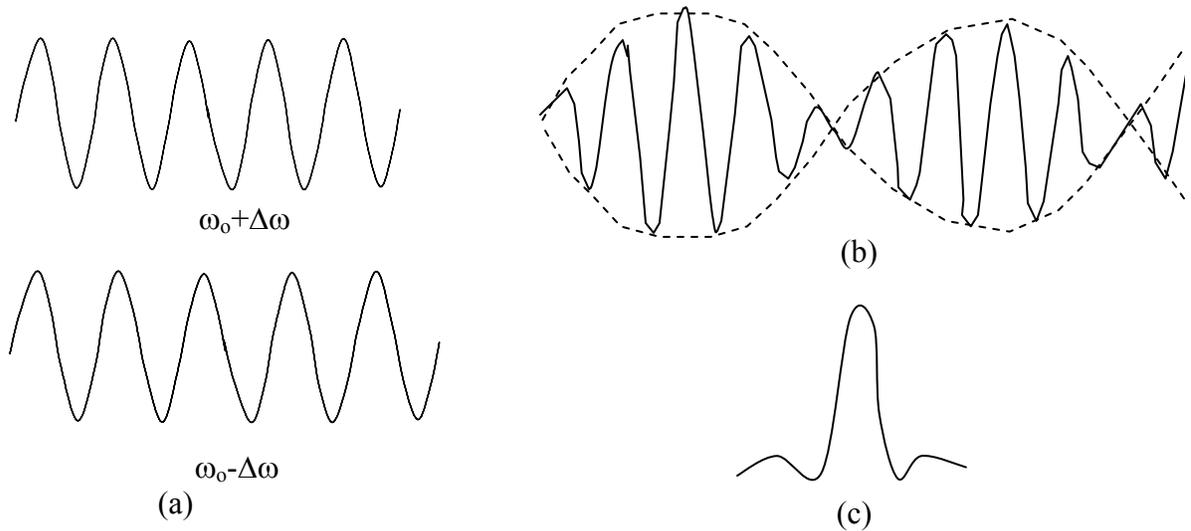
Such a finite-time harmonic signal can be decomposed through Fourier series into the summation of truly plane waves with time extending from minus infinite to plus infinite, as shown in figure (b). The frequencies of these plane waves are centered around ω_0 and their amplitudes decay as frequency moves away from ω_0 , as illustrated in figure (c).

One can better understand these pictures by actually carrying out the Fourier expansion ($\sum a_n \sin \frac{2\pi n}{T} t$). Because each of the plane waves in such a series expansion is at a frequency slight different from the central frequency ω_0 , it also has a corresponding wavevector that is different from k_0 , as determined by the dispersion relation between $\omega(k)$. The subsequent propagation of the signal can be obtained from tracing the spatial evolution of all these Fourier components as a function of time.

For simplicity, let's consider that the signal is an electromagnetic wave. We pick up only two of the Fourier components and consider their superposition, one at frequency $\omega_0 - \Delta\omega/2$ and another at frequency $\omega_0 + \Delta\omega/2$ propagating along positive x-direction [figure (a)]. The superposition of these two waves gives the electric fields as

$$E_y(x,t) = a \exp\{-i[(\omega_0 - \Delta\omega)t - (k_0 - \Delta k)x]\} + a \exp\{-i[(\omega_0 + \Delta\omega)t - (k_0 + \Delta k)x]\}$$

$$= 2a \cos(\Delta\omega \cdot t - \Delta k \cdot x) \exp[-i(\omega_0 t - k_0 x)]$$



The electric field is schematically shown in above figures. There appears to be two waves, one is the carrier wave at central frequency ω_0 (term $\exp[-i(\omega_0 t - k_0 x)]$), another is the modulation of the carrier wave by a wave at frequency $\Delta\omega$ (the amplitude term $2a \cos(\Delta\omega \cdot t - \Delta k \cdot x)$ in $E_y(x,t)$). The latter one changes much slower compared with the former one.

Note: (1) In the product $\Psi \nabla \Psi^*$ for flux J , the exponential term $\exp[-i(\omega_0 t - k_0 x)]$ is cancelled out. The time influence only exists in $2a \cos(\Delta\omega \cdot t - \Delta k \cdot x)$. (2) We need to calculate the time-averaged Poynting vector $\frac{1}{2} \text{Re}(\vec{E} \times \vec{H}^*)$ to get the real energy flux. This yields another wave propagating at the speed

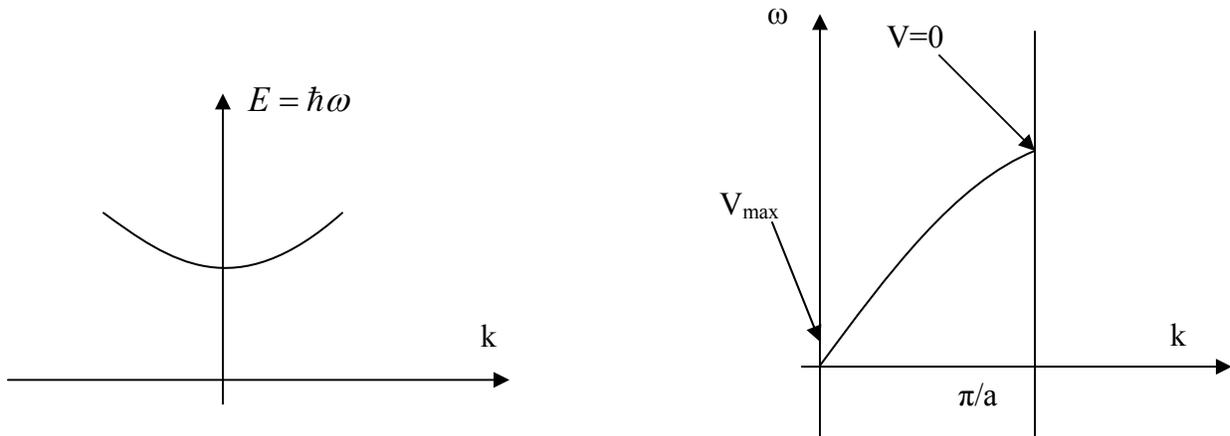
$$v_{g,x} = \frac{\Delta\omega}{\Delta k} = \frac{d\omega}{dk},$$

which means that the energy is propagating at the speed of $v_{g,x}$ rather than the phase velocity. We can also find this velocity from $\Delta\omega \bullet t - \Delta k \bullet x = 0$ in $E_y(x,t)$. This v_g is called the group velocity. In the more general case of the existence of a spectrum of frequencies, the superposition of waves leads to a narrow **wave packet** as sketched in above figure (c). The group velocity can be calculated from

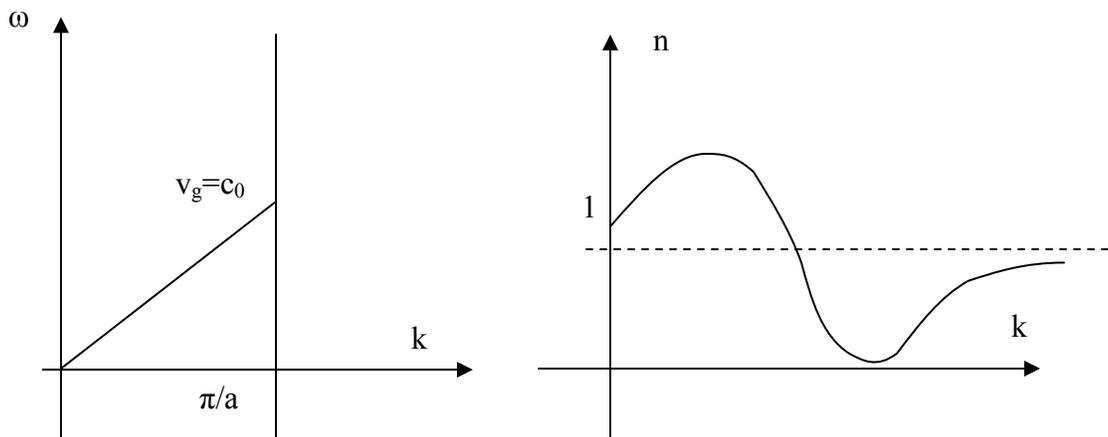
$$\mathbf{v}_g = \nabla_{\mathbf{k}} \omega = \frac{\partial \omega}{\partial k_x} \hat{k}_x + \frac{\partial \omega}{\partial k_y} \hat{k}_y + \frac{\partial \omega}{\partial k_z} \hat{k}_z.$$

Discussions

(1) In the following left figure, electrons in the conduction band have different group velocities (gradient in all figures) at different wave vectors. This is also the case for phonons, as shown in the right figure.

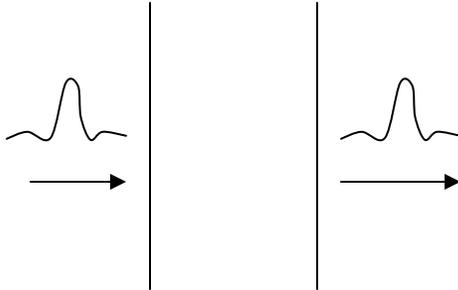


For photons, the group velocity is constant (the following left figure).



(2) In most cases, the energy velocity is just the group velocity. However, for photons the group velocity can be larger than the speed of light in special cases. We can appreciate this from the calculation of Poynting vector, where we assumed that $\Delta\omega$ is much smaller

than ω_0 . In the case of a very large variation in the dispersion relation, the group velocity no longer represents the energy velocity. Suppose we have a material with refractive index $n(\omega)$ changing rapidly with wavelength, also drawn as a function of k in the above right figure. The speed of light inside the material is $c = \frac{c_0}{n(\omega)}$. In a region, the real part of the refractive index is less than 1 and the phase velocity is larger than the speed of light. The front of the wave packet also moves with a speed larger than c_0 . This paradox is explained by noticing the majority energy still lies in the part behind the wave front. Therefore, the real energy speed is smaller than the speed of light.



Brillouin wrote his thesis on the group, signal velocities. For personal interests, you may want to read it for better understanding.

(3) The group velocity of electrons is velocity at which the electron wave packets moves in free space and inside a crystal. The energy dispersion of a free electron is

$$E = \frac{(\hbar\mathbf{k})^2}{2m}$$

and thus the phase and group velocity are, respectively,

$$\mathbf{v} = \frac{E/\hbar}{\mathbf{k}} = \frac{\hbar\mathbf{k}}{2m} \quad \text{and} \quad \mathbf{v}_g = \frac{\partial(E/\hbar)}{\partial\mathbf{k}} = \frac{\hbar\mathbf{k}}{m}$$

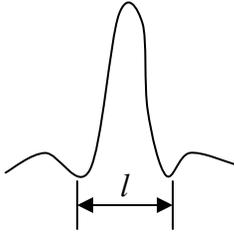
Clearly, the group velocity is consistent with the de Broglie relation $\mathbf{p} = \hbar\mathbf{k}$ and our classical relation $\mathbf{p} = m\mathbf{v}_g$, but not the phase velocity.

5.6.2 Loss of coherence

(1) Inelastic scattering

The scattering of electrons can be **elastic** in which the electrons merely change directions but have the same energy before and after the scattering, or **inelastic** in which both direction and energy of electrons are changed. The scattering of electrons by impurities and at the boundaries is elastic. The elastic scattering itself does not destroy the phase but the random locations of the impurities and the surface roughness, may create enough randomness in the phase such that the particle approach is approximately valid, or in other cases, the randomness can also create localization of the electron waves. The inelastic scattering, such as electron-phonon scattering, however, randomizes the phases because the location and the phase of the electron-phonon scattering change all the time.

In our previous discussion of transmissivity, we did not include the inelastic scattering and treat it as particle transport instead of using phase of wave functions.



For the above wave packet, the coherence length is defined as

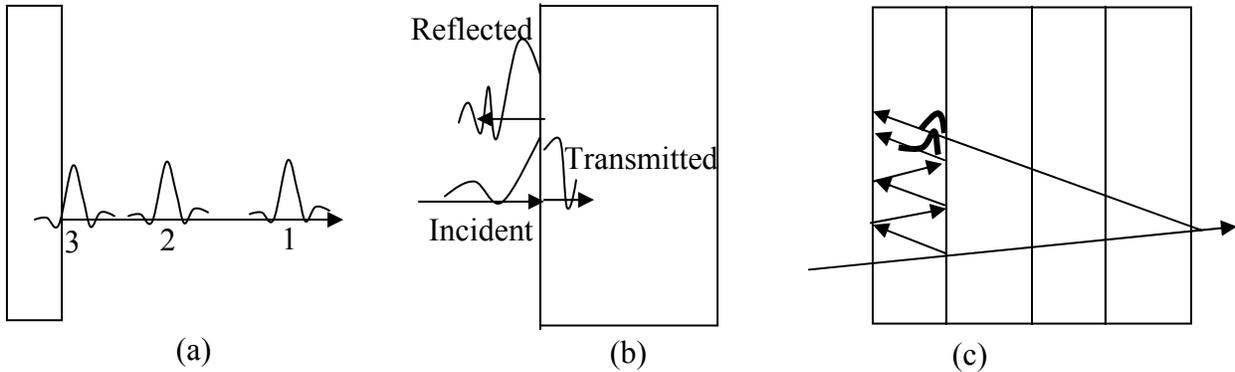
$$l \sim \frac{V}{\Delta \nu},$$

where $\Delta \nu$ is determined by $2\Delta\omega$ in the foregoing discussion. For lasers, $\Delta \nu$ is small and l can be on the magnitude of kilometers.

Now let us estimate the coherence length of blackbody radiation. The energy uncertainty of the individual radiation emitters (atoms, electrons, or molecules) is of the order of $k_B T$ due to the collision of the emitters with the reservoir, which also means that the effective bandwidth for thermal emission is $\Delta \nu = k_B T / h$. Using this effective bandwidth, one can estimate that the coherence length $\sim hc / (k_B T) = \frac{3E8 * 6.6E-34}{1.38E-23 * 300} = 3E-5$, which corresponds to $l_c T = 2167.8 \mu\text{m.K}$, very close to the Wien's displacement law.

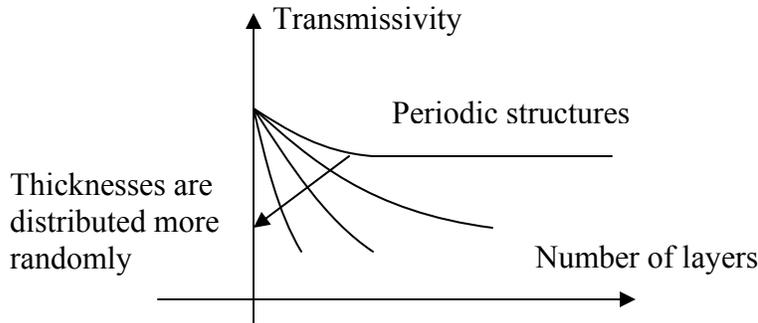
For electrons, $l_c = 100\text{-}1000 \text{ \AA}$, while it is $1\text{-}10 \text{ \AA}$ for phonons.

2) Spatial & temporal coherence



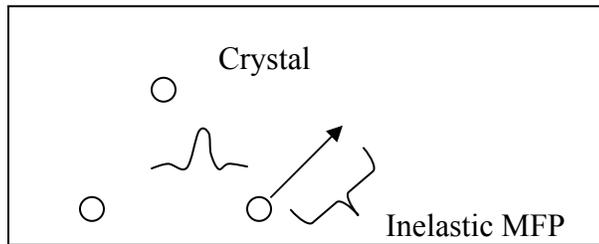
The coherence length as a measure of the wave packet size gives an indication whether the phase information needs to be considered for the transport processes or not. If the size of the transport domain is much larger than the wave packets or the coherence length, then the wave packets can be treated as point-wise particles traveling through the domain, as shown in above figure (a). When a wave packet meets a perfect interface, however,

the wave packet will be reflected and refracted. The refracted wave packet has a fixed phase relationship with the incoming one and can thus interfere with the incoming wave packet [figure (b)]. This is the reason why we always use Fresnel formula---the wave solution of the Maxwell equations, to calculate the interface reflectivity and transmissivity of a perfect interface. In the case of multilayers, there is still overlap between the incoming and reflected wave packet even if the layer is very thick [figure (c)]. If we alter the thickness of each layer randomly, the overlap still exists. The transmissivity is drawn in the following figure.

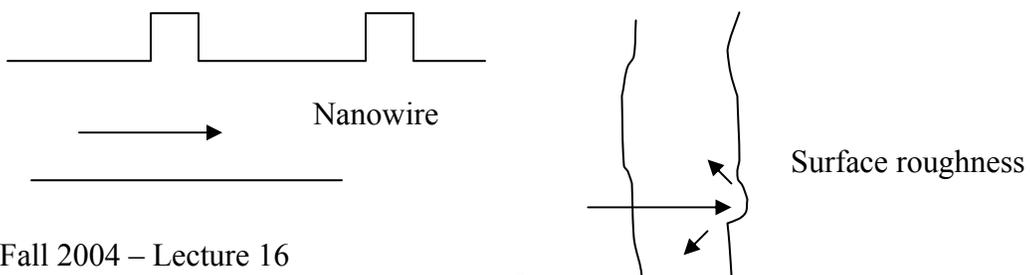


Note: Randomly distributed thickness will cause localization phenomena and reduce the transmissivity, which won Phillip Anderson a Nobel Prize in physics.

Consider scattering in the following crystal. The wave is scattered at random points due to inelastic scattering such as by phonons. In this case, we can ignore the phase information and the particle picture is valid.



When size effects appear, the characteristic length is normally less than the inelastic mean free path. This can be the case for nanowires. However, the particle treatment may be still valid due to the boundary scattering. Although boundary scattering is not random and most likely elastic, the roughness diffracts waves into all directions. The averaging of the diffracted waves usually leads to results more closely to particle treatment than these based ideal interfaces. For phonons in a film with surface roughness, we can argue in the same way.

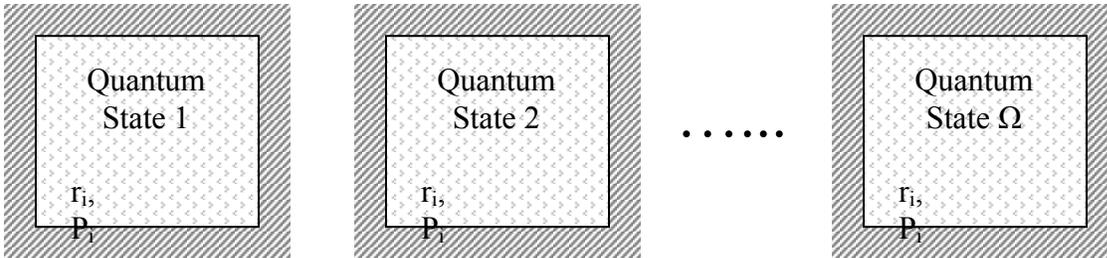


Chapter 6 Particle Description of Transport Process

6.1 Liouville equation and Boltzmann equation

In chapter 4, we studied systems at equilibrium and developed the equilibrium distribution functions $f(E, T, \mu)$ (Fermi-Dirac, Bose-Einstein, and Boltzmann distributions). The distribution function for a quantum state at equilibrium is a function of the energy of the quantum state, the system temperature and chemical potential. When the system is not at equilibrium, these distribution functions are no longer applicable. In the statistical description, we use nonequilibrium distribution functions, which depend not only on the energy and temperature of the system, but also on positions and other variables. We will develop in this chapter the governing equations for the nonequilibrium distribution functions.

In the following figure, we show an ensemble, i.e., a collection of quantum states of a N-particle system.



Now each particle can be described by the generalized coordinate \mathbf{r} and momentum \mathbf{p} . For example, the generalized coordinates of a diatomic molecule, \mathbf{r}_1 , include the position (x_1, y_1, z_1) , the vibrational coordinate (the separation between the two atoms, Δx_1), the rotational coordinates (polar and azimuthal angles, θ_1 and ϕ_1), and the generalized momentum, \mathbf{p}_1 , includes the translational $(mv_{x1}, mv_{y1}, mv_{z1})$, vibrational momentum proportional to the relative velocity of the two atoms $(m d\Delta x_1/dt)$, and rotational momentum (two angular momentum of rotation). We assume here that there are m -degrees of freedom in space, i.e., m generalized spatial coordinates, and m -degrees of freedom in momentum for each particle. The number of the degree of freedom of the whole system is $2n = 2m \times N$. These $2n$ variables form a $2n$ -dimensional space. The system at any instant can be described as one point in such a space. This space is called a phase space. The time evolution of the system, i.e., the time history of all the particles in the system, traces one line in such a $2n$ -dimensional **phase space**, which we will call the flow line as in fluid mechanics.

The number of systems, is

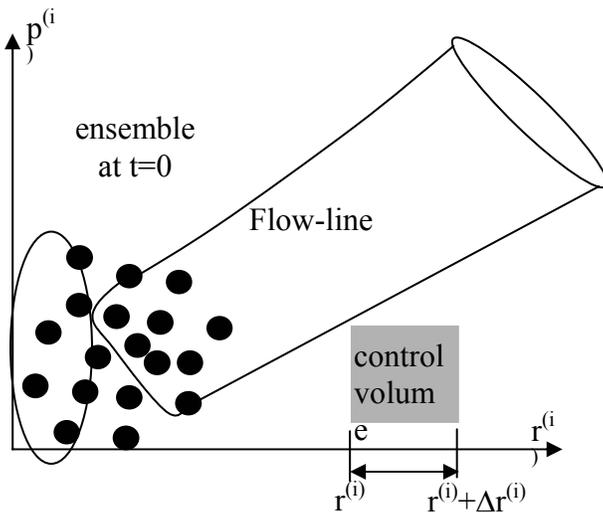
$$\text{No. Systems} = f^{(N)}(\mathbf{t}, \mathbf{r}^{(n)}, \mathbf{p}^{(n)}) \Delta \mathbf{r}^{(n)} \Delta \mathbf{p}^{(n)}$$

in a small volume of the phase space, $\Delta \mathbf{r}^{(n)} \Delta \mathbf{p}^{(n)}$, where $\Delta \mathbf{r}^{(n)} = \Delta r_1 \Delta r_2 \dots \Delta r_N = \Delta r^{(1)} \Delta r^{(2)} \dots \Delta r^{(n)}$ and $\Delta \mathbf{p}^{(n)} = \Delta p_1 \Delta p_2 \dots \Delta p_N = \Delta p^{(1)} \Delta p^{(2)} \dots \Delta p^{(n)}$.

The time evolution of $f^{(N)}(t, \mathbf{r}^{(n)}, \mathbf{p}^{(n)})$ in the phase space is governed by the **Liouville equation**, which can be derived based on the fact that the traces of systems in the ensemble do not intersect. Consider a tube formed by the traces of a set of points (a subset of systems in the ensemble) as shown in the following figure. Since the flow lines do not intersect, the points in the phase space are conserved. We want to derive an equation for the distribution function f based on this conservation requirement. Recall that in fluid mechanics or heat transfer, we often use the control volume method rather than tracing the trajectory of individual fluid particles. We could do the same for the points in phase space and exam a small control volume in phase space as shown in the figure. The rate of points flow into the control volume should equal to the rate of change inside the control volume. This finally leads to

$$\frac{\partial f^{(N)}}{\partial t} + \sum_{i=1}^n \dot{r}^{(i)} \times \frac{\partial f^{(N)}}{\partial r^{(i)}} + \sum_{i=1}^n \dot{p}^{(i)} \times \frac{\partial f^{(N)}}{\partial p^{(i)}} = 0$$

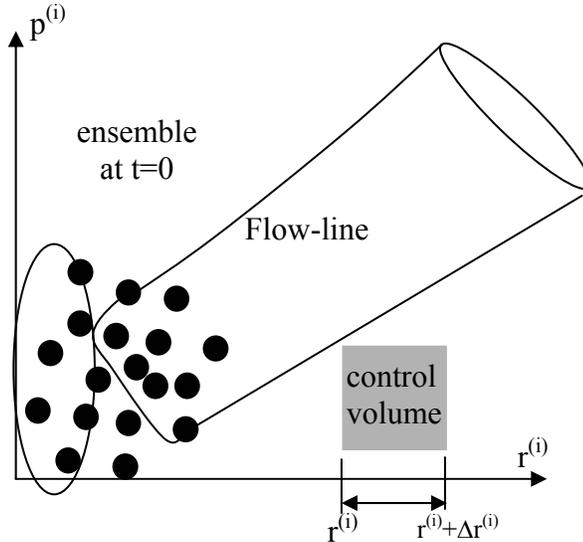
which is called the Liouville equation.



Note: The Boltzmann equation reduces variables from the Liouville equation, while the linear response theory focuses on small perturbation. They are both simplified form of the Liouville equation.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 17

In last lecture, we discussed the N-particle distribution function. We have obtained
 No. Systems = $f^{(N)}(t, \mathbf{r}^{(n)}, \mathbf{p}^{(n)}) \Delta \mathbf{r}^{(n)} \Delta \mathbf{p}^{(n)}$
 in a small volume of the phase space, $\Delta \mathbf{r}^{(n)} \Delta \mathbf{p}^{(n)}$,
 where $\Delta \mathbf{r}^{(n)} = \Delta r_1 \Delta r_2 \dots \Delta r_N = \Delta r^{(1)} \Delta r^{(2)} \dots \Delta r^{(n)}$ and $\Delta \mathbf{p}^{(n)} = \Delta p_1 \Delta p_2 \dots \Delta p_N = \Delta p^{(1)} \Delta p^{(2)} \dots \Delta p^{(n)}$.



The time evolution of $f^{(N)}(t, \mathbf{r}^{(n)}, \mathbf{p}^{(n)})$ in the phase space is governed by the **Liouville equation**,

$$\frac{\partial f^{(N)}}{\partial t} + \sum_{i=1}^n \dot{r}^{(i)} \times \frac{\partial f^{(N)}}{\partial r^{(i)}} + \sum_{i=1}^n \dot{p}^{(i)} \times \frac{\partial f^{(N)}}{\partial p^{(i)}} = 0$$

which can be derived based on the fact that the traces of systems in the ensemble do not intersect.

Note: The number of the degree of freedom in the phase space is normally very big. For 3D cases, it is $6N_A = 6 \times 6.02E23$.

The Liouville equation involves a huge number of variables, which makes its impractical in terms of the boundary and initial conditions, as well as analytical and numerical solutions. One way to simplify the Liouville equation is to consider one particle in a system. This is a representative particle having coordinate \mathbf{r}_1 and momentum \mathbf{p}_1 , each of the vectors has m components, i.e., m -degrees of freedom. We introduce a **one-particle distribution function** by averaging the N-particle distribution function over the rest (N-1) particles in the system,

$$f^{(1)}(t, \mathbf{r}_1, \mathbf{p}_1) = \frac{N!}{(N-1)!} \int \dots \int f^{(N)}(t, \mathbf{r}^{(n)}, \mathbf{p}^{(n)}) d\mathbf{r}_2 \dots d\mathbf{r}_N d\mathbf{p}_2 \dots d\mathbf{p}_N.$$

For simplicity in notation, we will drop the subscript 1 and understand (\mathbf{r}, \mathbf{p}) as the coordinates and momenta of one particle. Since $f^{(N)}(t, \mathbf{r}^{(n)}, \mathbf{p}^{(n)})$ represents the number density of systems having generalized coordinates $(\mathbf{r}^{(n)}, \mathbf{p}^{(n)})$ in the ensemble, the one particle distribution function represents number density of systems having (\mathbf{r}, \mathbf{p}) , $f(t, \mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p}$ = number of systems in $d^3\mathbf{r} d^3\mathbf{p}$.

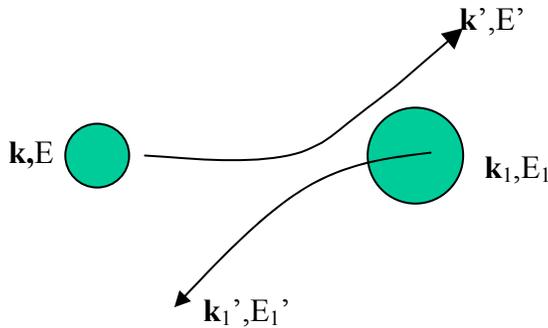
With the introduction of the averaging method to obtain the one-particle distribution function, one can start from the Liouville equation and carry out the averaging over the space and momentum coordinates of the other (N-1) particles. This procedure leads to

$$\frac{\partial f}{\partial t} + \frac{d\mathbf{r}}{dt} \cdot \nabla_{\mathbf{r}} f + \frac{d\mathbf{p}}{dt} \cdot \nabla_{\mathbf{p}} f = \left(\frac{\partial f}{\partial t} \right)_c$$

where the subscripts $(\mathbf{r}$ and $\mathbf{p})$ in the gradient operators represent the variables of the gradient. The scattering term $\left(\frac{\partial f}{\partial t} \right)_c$ will be discussed in details later in this lecture. The

above equation is the **Boltzmann equation** or Boltzmann transport equation.

Note: The derivative $\frac{d\mathbf{r}}{dt}$ has the meaning of velocity, while $\frac{d\mathbf{p}}{dt}$ denotes force.



Consider the collision process between two particles as shown in the above figure. After the collision, the energy and the velocity of each particle may change. Clearly, the collision is a time-dependent process. The rigorous way of dealing the collision process is to solve the corresponding time-dependent Schrödinger equation for the combined system made of both particles. This is, however, usually very complicated and not practical. A simpler way to treat the collision is to use the perturbation method. This method considers the time-dependent interaction between the two particles as a small perturbation in energy, $H'(\mathbf{r}, t)$, from the original steady-state, non-interacting energy H_0 of the two particles, such that the total system energy is

$$H = H_0(r) + H'(\mathbf{r}, t),$$

where we have learned that the Hamiltonian H_0 is

$$H_0 = \frac{\hat{p}^2}{2m} + U = \frac{1}{2m} (-i\hbar\nabla)^2 + U = \frac{-\hbar^2}{2m} \nabla^2 + U$$

in the Schrödinger equation $H_0\Psi = E\Psi$.

By treating H' as a small perturbation to unperturbed Hamiltonian H_0 , the solution of the Schrödinger equation for the new H can be obtained through perturbation method and expressed in terms of the wave functions Ψ of the unperturbed two-particle system with a Hamiltonian H_0 . Using the perturbation solution, one can calculate the probability for the system jumping from one quantum state Ψ_i to another quantum state Ψ_f , both are accessible quantum states of the original two-particle systems. This rate of this probability is the transition rate,

$$\begin{aligned} W_{if}^f(\mathbf{k}, \mathbf{k}', \mathbf{k}_1, \mathbf{k}_1') &= \frac{2\pi}{\hbar} \left[\int \Psi_f^* H' \Psi_i d^3\mathbf{r} \right]^2 \delta(E_f - E_i) \\ &= \frac{2\pi}{\hbar} \left| \langle i | H' | f \rangle \right|^2 \delta(E_f - E_i) \\ &= \frac{2\pi}{\hbar} M_{if}^2 \delta(E_f - E_i) \end{aligned}$$

where $d^3\mathbf{r} = dx dy dz$ means integration over the whole volume of the system and

$$M_{if} \equiv \langle i | H' | f \rangle \equiv \int \Psi_f^* H' \Psi_i d^3\mathbf{r}$$

is called the scattering matrix. The Kronecker delta function, $\delta(E_f - E_i)$, implies that the energy must be conserved in the process. This equation is one form of the often-referred **Fermi golden rule**.

The scattering term in the Boltzmann equation is the net gain of particles in one quantum state. This net gain consists of two components: one is the increase of the number of particles due to scattering from other quantum states into the quantum state under consideration. The other is the decrease of the number of particles due to scattering from the current quantum states to other quantum states. We again take two-particle scattering process as an example. The initial wave vector of the particle is \mathbf{k} and it collides with a particle with a wavevector \mathbf{k}_1 . The corresponding distribution functions for the two particles are $f(\mathbf{r}, \mathbf{k}, t)$ and $f(\mathbf{r}, \mathbf{k}_1, t)$. After scattering, the momentum of the two particles are \mathbf{k}' and \mathbf{k}_1' and their distribution functions are $f(\mathbf{r}', \mathbf{k}', t)$ and $f(\mathbf{r}', \mathbf{k}_1', t)$, respectively. The scattering term for the particle at state \mathbf{k} can be expressed as

$$\begin{aligned} \left(\frac{\partial f}{\partial t} \right)_c &= - \int f(\mathbf{r}, \mathbf{k}, t) f(\mathbf{r}, \mathbf{k}_1, t) W(\mathbf{k}, \mathbf{k}_1 \rightarrow \mathbf{k}', \mathbf{k}_1') d^3\mathbf{k}_1 d^3\mathbf{k}' d^3\mathbf{k}_1' \\ &\quad + \int f(\mathbf{r}, \mathbf{k}', t) f(\mathbf{r}, \mathbf{k}_1', t) W(\mathbf{k}', \mathbf{k}_1' \rightarrow \mathbf{k}, \mathbf{k}_1) d^3\mathbf{k}_1 d^3\mathbf{k}' d^3\mathbf{k}_1', \end{aligned}$$

where the first term on the right hand side represents the rate of particles being scattered out of quantum states determined by \mathbf{k} and \mathbf{k}_1 , and the second term represents rate of particles scattered into the quantum state.

When discussing the interfacial thermal conductance in Lecture 14, we have obtained $q_{1 \rightarrow 2}(T_{e2}) = q_{2 \rightarrow 1}(T_{e2})$ by considering the equilibrium status $q_{12} = 0$ and $T_{e2} = T_{e1}$. Based on

a similar argument, the above equation can be simplified by noticing in the equilibrium status

$$W(\mathbf{k}, \mathbf{k}_1 \rightarrow \mathbf{k}', \mathbf{k}_1') = W(\mathbf{k}', \mathbf{k}_1' \rightarrow \mathbf{k}, \mathbf{k}_1).$$

This leads to

$$\left(\frac{\partial f}{\partial t}\right)_c = -\frac{V^3}{(2\pi)^9} \int W(\mathbf{k}, \mathbf{k}_1 \rightarrow \mathbf{k}', \mathbf{k}_1') [f(\mathbf{r}, \mathbf{k}, t) f(\mathbf{r}, \mathbf{k}_1, t) - f(\mathbf{r}, \mathbf{k}', t) f(\mathbf{r}, \mathbf{k}_1', t)] d^3\mathbf{k}_1 d^3\mathbf{k}' d^3\mathbf{k}_1'$$

where the factor $\frac{V^3}{(2\pi)^9}$ appears in the conversion from summation over wavevector into integration over the phase space.

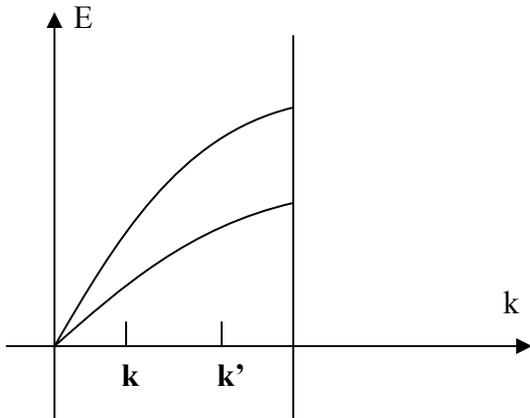
The integral-differential Boltzmann equation is very difficult to solve in general. Most solutions rely on a drastic simplification of the scattering integral by the relaxation time approximation,

$$\left(\frac{\partial f}{\partial t}\right)_c = -\frac{f - f_0(T, E, \mu)}{\tau(\mathbf{r}, \mathbf{k})},$$

where $\tau(\mathbf{r}, \mathbf{k})$ is the **relaxation time**, and f_0 represents the equilibrium distribution of the carriers, such as the Boltzmann, the Fermi-Dirac, and the Bose-Einstein distributions given in chapter 4. Under the relaxation time approximation, the Boltzmann equation becomes

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f = -\frac{f - f_0}{\tau}.$$

Note: In fluid mechanics, we have a counterpart of the Boltzmann equation, which is called the Krook equation.



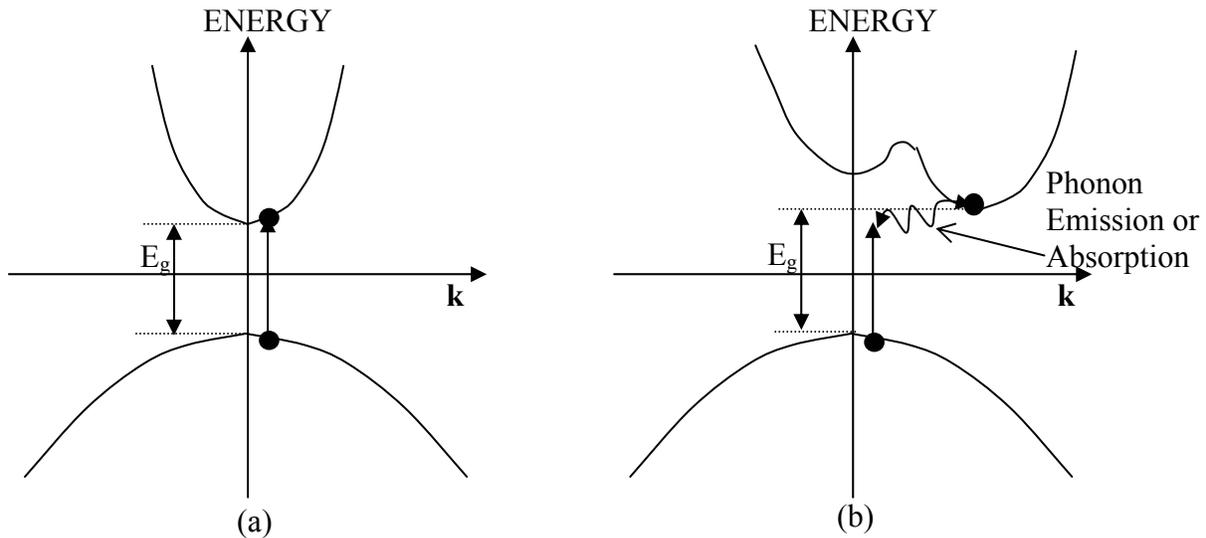
Now let us consider the general rules of the phonon-phonon scattering. First we have energy balance as

$$E_f = E_i \Rightarrow E(\mathbf{k}_1) + E(\mathbf{k}) = E(\mathbf{k}_1') + E(\mathbf{k}').$$

Secondly we have momentum conservation

$$\mathbf{G} + \mathbf{k}_1 + \mathbf{k} = \mathbf{k}_1' + \mathbf{k}',$$

where the vector \mathbf{G} accounts for the momentum imbalance between the initial and final states when $|\mathbf{k}_i + \mathbf{k}| > \pi/a$. We have learned that $|\mathbf{k}_i + \mathbf{k}| > \pi/a$ (out of the first Brillouin zone) is meaningless and must be flipped over into the first Brillouin zone. This momentum difference is compensated by \mathbf{G} . We call this process umklapp process. In other case, $\mathbf{G}=0$ and it is called a normal process.

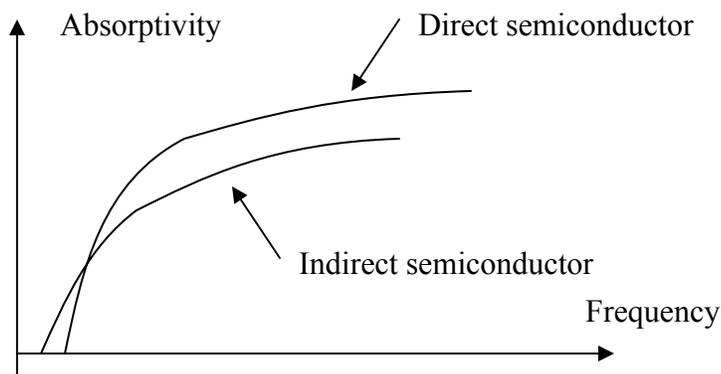


The absorption of photons by free electrons in metals and in semiconductors must be accompanied by emission or absorption of phonons because of momentum conservation. In the above figures, figure (a) shows that an electron absorbs a photon and is pumped to a different band (direct semiconductor). We still have

$$E_i + \hbar\omega = E_f \text{ (energy conservation)}$$

$$\mathbf{k}_i + \frac{\hbar}{\lambda} = \mathbf{k}_f \text{ (momentum conservation),}$$

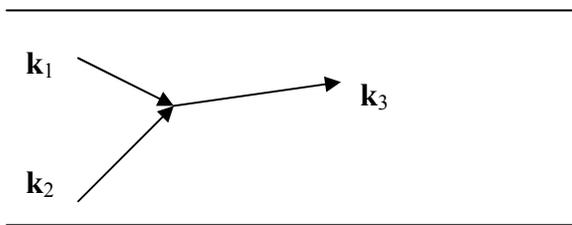
where $\frac{\hbar}{\lambda}$ is the photon momentum and can be proved to be negligible for optical lights ($\lambda \sim 0.5 \mu m$). Therefore, the arrow points upward almost vertically.



For silicon (indirect semiconductor), we have the situation drawn in figure (b). In this case, the absorption of a photon cannot happen at the band edge without the absorption or emission of phonons because photon-electron interaction alone does not satisfy the wave vector conservation requirements. The participation of phonons in the process reduces the absorption or emission of photons, as shown in the above figure.

Generally speaking, electrons will collide with electrons, phonons, and impurities in the crystals. For a normal process in which two phonons merge into one (shown in the figure below), both the energy and momentum are conserved and will not create any thermal resistance to the heat conduction. In fact, the finite thermal conductivity of a crystal results from the random umklapp process with compensating \mathbf{G} . It has the relaxation time

$$\tau_u^{-1} = B e^{-\theta_D / bT} T^3 \omega^2$$



For impurity scattering, we have the familiar Rayleigh law

$$\tau_I^{-1} = A \omega^4$$

Note the scattering possibility is proportional to the reciprocal of relaxation time. This equation suggests that the scattering is enhanced for increased ω .

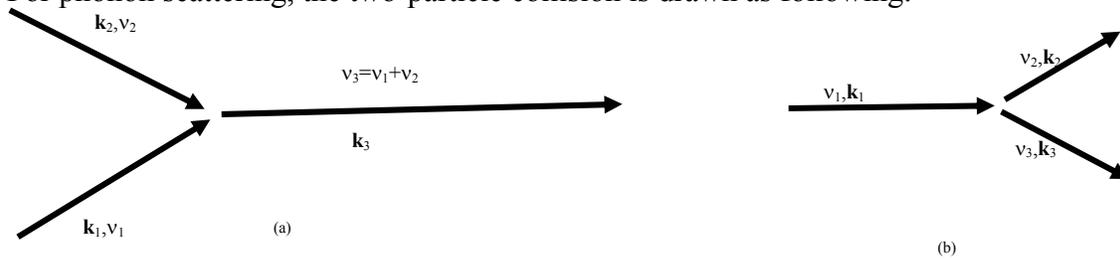
Note: (1) Normally we assume the size of defects $d \ll \lambda$. (2) The blue sky is the result of impurity scattering that shields off the high frequency lights.

The total relaxation time is obtained by combining the expressions for individual relaxation processes according to the Mathiessen rule

$$\frac{1}{\tau} = \sum \frac{1}{\tau_j}$$

which is used in the expression of $\frac{f - f_o}{\tau}$.

For phonon scattering, the two-particle collision is drawn as following.



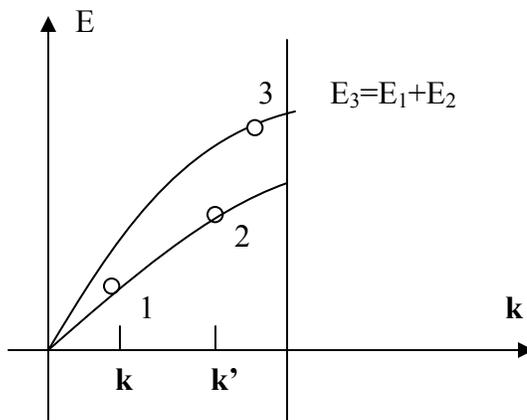
For the merging of two phonons into one, the energy conservation gives

$$h \nu_1 + h \nu_2 = h \nu_3$$

and a similar equation can be written for the process that one phonon splits into two. The momentum conservation during the three-phonon interaction processes takes a special form. For the phonon merging process, the momentum conservation can be written as

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{G}$$

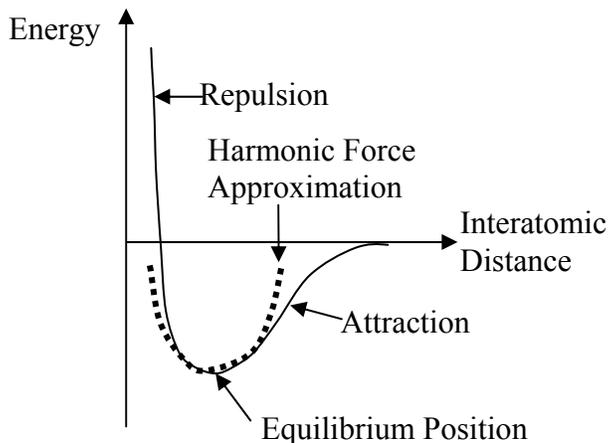
where the reciprocal lattice vector \mathbf{G} can be zero or a linear combination of the reciprocal lattice vectors. If $(\mathbf{k}_1 + \mathbf{k}_2)$ is smaller than the first Brillouin zone wavevector, \mathbf{G} equals $\mathbf{0}$, otherwise, $\mathbf{G} \neq \mathbf{0}$. This comes from the requirement that the phonon wavelength cannot be smaller than the lattice constant as we discussed in chapter 3. The $\mathbf{G} = \mathbf{0}$ phonon scattering process is called the **normal process** and the $\mathbf{G} \neq \mathbf{0}$ is the **umklapp process**, as mentioned before.



Note: In the perturbation method, we consider the time-dependent interaction between the two particles as a small perturbation in energy, $H'(\mathbf{r}, t)$. When we use the harmonic oscillator approximation for the actual interatomic potentials, the higher order term $O[(x' - x_0)^3]$ in

$$U(x') = U(x_0) + \frac{1}{2} \left[\frac{d^2 U}{dx'^2} \right]_{x'=x_0} (x' - x_0)^2 + O[(x' - x_0)^3]$$

can be considered as the perturbation from the harmonic potential. In the umklapp process, the randomly distributed vector \mathbf{G} plays a similar role as $O[(x' - x_0)^3]$.



When the particles have a nonzero average velocity \mathbf{u} , the following **displaced Maxwell velocity distribution** is often used for f_0

$$f_0(\mathbf{v}) = n \left(\frac{m}{2\pi\kappa_B T} \right)^{3/2} e^{-m[(\mathbf{v}-\mathbf{u})^2]/2\kappa_B T},$$

where \mathbf{u} can change at different location in the space.

If the solution of the Boltzmann equation is known for a problem, we can calculate the volume-average of any microscopic property X of the particle from,

$$\langle X(\mathbf{r}) \rangle = \frac{1}{V} \sum_{\mathbf{k}, s} X(\mathbf{r}, \mathbf{k}) f,$$

where the summation can be conducted over wave vector \mathbf{k} or velocity \mathbf{v} .

Assumption of the Boltzmann equation

- (1) The Boltzmann equation is only applicable to diluted systems, such as phonons, electrons, molecules, and photons. It can never be used for liquids.
- (2) Consider the collision of two particles. Before and after the collision, the distribution functions of the one of the two particles are independent of the coordinates and momentum of the other particle. This is the so called the **molecular chaos** assumption. Additionally, the particles have no memory of its history and its final state is unaffected by its very beginning status.
- (3) The Boltzmann equation cannot include explicitly the wave effects such as interference and tunneling. We use the particle picture for all particles though photons have larger wavelength and is more like a wave. The wave theory is partially included for some problems. For example, we have solved the wavefunctions of electrons in deriving the band structure in crystals. The wavelength is far larger than the interatomic distance.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 18

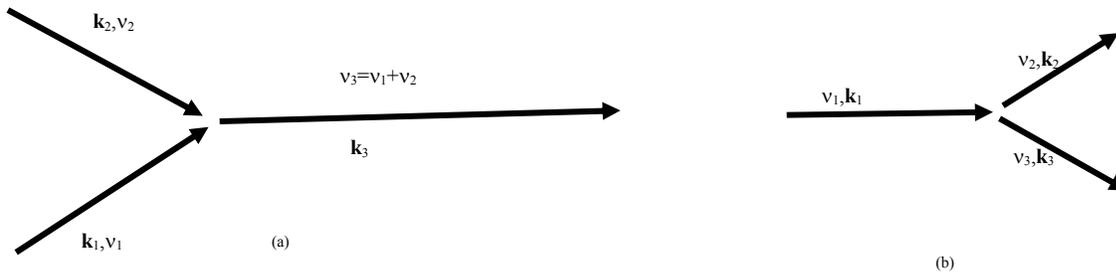
In last lecture, we deal with the Boltzmann transport equation

$$\frac{\partial f}{\partial t} + \frac{d\mathbf{r}}{dt} \cdot \nabla_{\mathbf{r}} f + \frac{d\mathbf{p}}{dt} \cdot \nabla_{\mathbf{p}} f = \left(\frac{\partial f}{\partial t} \right)_s$$

where the subscripts (**r** and **p**) in the gradient operators represent the variables of the gradient. The scattering term $\left(\frac{\partial f}{\partial t} \right)_s = -\frac{f - f_0}{\tau}$ was discussed based on two particle interactions. Here $\tau(\omega, \mathbf{k})$ is the relaxation time. For equilibrium distribution f_0 , we have

$$f_0 = \begin{cases} \frac{1}{e^{\frac{\hbar\omega}{k_B T}} - 1} & \text{Bose-Einstein distribution (phonon)} \\ \frac{1}{e^{\frac{E-\mu}{k_B T}} + 1} & \text{Fermi-Dirac distribution (electron)} \\ n \left(\frac{m}{2\pi k_B T} \right)^{3/2} e^{-\frac{m|\mathbf{v}-\mathbf{u}|^2}{2k_B T}} & \text{Displaced Maxwell velocity distribution (molecules)} \end{cases}$$

Note: The relaxation time is due to combined factors and can be evaluated numerically by adding all possible influence together. This idea can be used in calculating the band structure by the Boltzmann transport equation.



For two-particle interactions (above figure (a)), we have

$$\hbar\omega_1 + \hbar\omega_2 = \hbar\omega_3 \quad (\text{energy conservation}),$$

$$\mathbf{G} + \mathbf{k}_3 = \mathbf{k}_1 + \mathbf{k}_2 \quad (\text{momentum conservation}),$$

where zero **G** corresponds to normal process, otherwise it is umklapp scattering.

Generally speaking, electrons will collide with electrons, phonons, and impurities in the crystals. The electron-phonon scattering causes the electrical resistance.

Note: At low temperatures, phonon has low energy and the electron-phonon scattering is negligible. Therefore, the impurity-electron scattering is the main cause of electrical resistance.

For gas molecules, in Lecture 2 we have derived the collision obeys

$$\Lambda = \frac{1}{\sqrt{2n\pi D^2}}, \quad \Lambda = \bar{v}\tau.$$

Note: (1) To simplify, most time we view $\tau(\omega, \mathbf{k})$ as $\tau(\omega)$. (2) The time $\tau(\omega, \mathbf{k})$ is only applicable to elastic scattering (see chapter 8), such as electron-electron scattering (energy conserved). It is not accurate for electron-phonon scattering. In this situation, we have $\left(\frac{\partial f}{\partial t}\right)_s = -\frac{f - f_0}{\tau} + g(T_e - T_p)$. (3) In semiconductor devices, T_e can be thousands of degrees, while T_p is only hundreds degrees. The two temperatures differ a lot.

We can understand the meaning of τ easily by neglecting the spatial non-uniformity of the distribution function. The Boltzmann transport equation becomes

$$\frac{\partial f}{\partial t} = -\frac{f - f_0}{\tau}$$

and thus

$$f - f_0 = C e^{-t/\tau}$$

So the relaxation time is a measure of how long it takes for a nonequilibrium system to relax back to an equilibrium distribution.

Consider the Boltzmann equation under the relaxation time approximation. Let's introduce a deviation function g ,

$$g = f - f_0$$

and write the Boltzmann equation as

$$\frac{\partial g}{\partial t} + \frac{\partial f_0}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 + \mathbf{v} \cdot \nabla_{\mathbf{r}} g + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_0 + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} g = -\frac{g}{\tau}$$

Note: $\nabla_{\mathbf{p}} f = \frac{1}{\hbar} \nabla_{\mathbf{k}} f = \frac{1}{m} \nabla_{\mathbf{v}} f$. We can also express $f(t, \mathbf{r}, \mathbf{p})$ as $f(t, \mathbf{r}, \mathbf{v})$ or $f(t, \mathbf{r}, \mathbf{k})$ for convenience.

All the diffusion laws can be obtained under the following assumptions (1) the transient terms are negligible, (2) the gradient of g is much smaller than the gradient of f_0 , and similarly, g is much smaller than f_0 . Under these assumptions, the above equation becomes

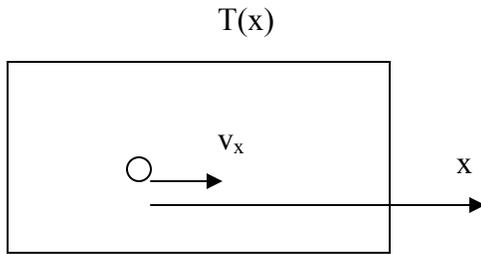
$$g = -\tau \left(\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_0 \right)$$

or

$$f = f_o - \tau \left(\mathbf{v} \cdot \nabla_{\mathbf{r}} f_o + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_o \right).$$

From the distribution function changing in long period compared with the relaxation time, we can calculate the flux of various quantities of interests (charge, momentum, and energy).

Fourier Law



For simplicity, we consider a temperature gradient along the x-direction without loss of generality. We can calculate the heat flux from

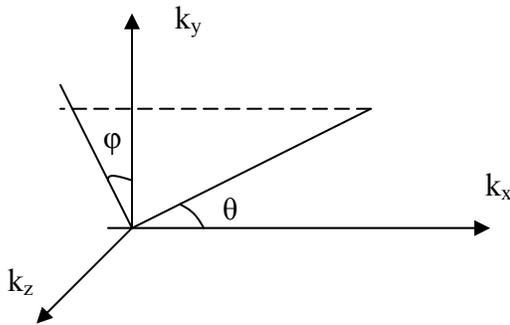
$$J_{qx}(x) = \sum_p \left[\frac{1}{V} \sum_{k_{x1}=-\infty}^{\infty} \sum_{k_{y1}=-\infty}^{\infty} \sum_{k_{z1}=-\infty}^{\infty} v_x \hbar \omega(\omega) f \right]$$

where p represents the summation over all polarizations. It is interesting to compare this expression with what we used in deriving the Landauer formalism. In that equation, we are considering only the heat flux going from point 1 to point 2 and there also exists a reverse heat flux from point 2 to point 1. Here we are considering the net heat flux at any constant x-plane inside the domain.

As in previous lectures, we can convert the summation into integral and further change into spherical coordinates

$$J_{qx}(x) = (1/V) \sum_p \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v_x \hbar \omega f dk_x dk_y dk_z / (2\pi/L)^3$$

$$= \int_0^{\omega_{\max}} d\omega \left[\int_0^{2\pi} \left\{ \int_0^{\pi} v \cos \theta \hbar \omega f \frac{D(\omega)}{4\pi} \sin \theta d\theta \right\} d\varphi \right]$$



Note: Here θ varies from 0 to π because both positive and negative v_x are considered.

With $f = f_o - \tau \left(\mathbf{v} \cdot \nabla_{\mathbf{r}} f_o + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_o \right)$, we obtain

$$\begin{aligned}
 J_{qx}(x) &= \int_0^{\omega_{\max}} d\omega \left[\int_0^{2\pi} \left\{ \int_0^{\pi} v \cos \theta \hbar \omega \left[f_o - \tau \frac{df_o}{dT} \frac{dT}{dx} v \cos \theta \right] \frac{D(\omega)}{4\pi} \sin \theta d\theta \right\} d\varphi \right] \\
 &= -\frac{1}{2} \frac{dT}{dx} \int_0^{\omega_{\max}} d\omega \left\{ \int_0^{\pi} \tau v^2 \sin \theta \cos^2 \theta \times \hbar \omega D(\omega) \frac{df_o}{dT} d\theta \right\} \\
 &= -\frac{1}{3} \frac{dT}{dx} \int_0^{\omega_{\max}} d\omega \left\{ \int_0^{\pi} \tau v^2 \frac{d(\hbar \omega D(\omega) f_o)}{dT} d\theta \right\} \\
 &= -\frac{1}{3} \frac{dT}{dx} \int_0^{\omega_{\max}} d\omega \left\{ \int_0^{\pi} \tau v^2 C(\omega) d\theta \right\}
 \end{aligned}$$

and

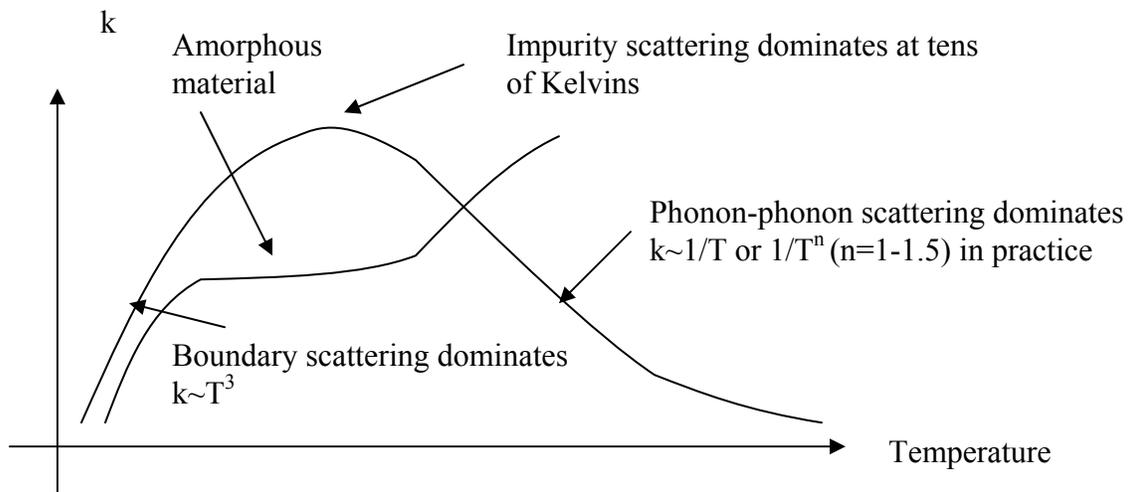
$$k = \frac{1}{3} \int \tau v^2 C d\omega \text{ in the Fourier law.}$$

Note: The equilibrium f_o term drops out in the integral, which is consistent with our expectation.

In the case that both τ and v (group velocity $v = d\omega/dk$) are independent of frequency, the above expression recover to the kinetic relation

$$k = \frac{1}{3} C v \Lambda.$$

The temperature-dependent thermal conductivity for a material is drawn in the following figure. For amorphous materials, two transition points exist.

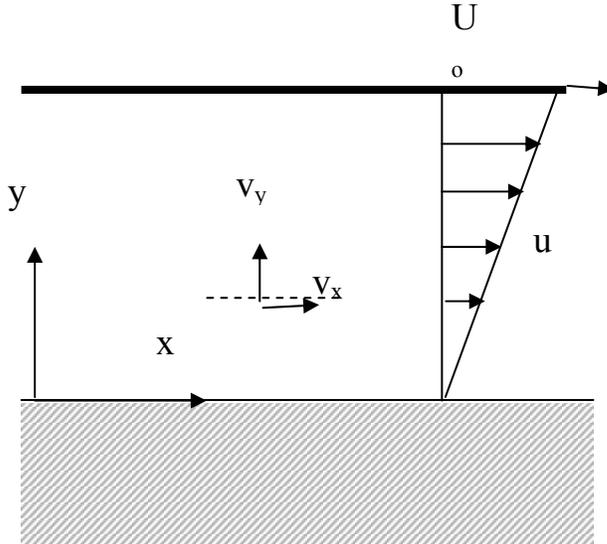


Newton's Shear Stress Law

In the following figure, a couette flow is shown. As an approximation, the following **displaced Maxwell velocity distribution** is often used for the probability of finding one particle having velocity \mathbf{v}

$$P(v_x, v_y, v_z) = \left(\frac{m}{2\pi\kappa_B T} \right)^{3/2} e^{-m \left[(v_x - u)^2 + v_y^2 + v_z^2 \right] / 2\kappa_B T}$$

where u is the average velocity along the x-direction (spatially changed).



Assuming that the number density of particles is n , the number density of particles having velocity \mathbf{v} is

$$f_0(v_x, v_y, v_z) = n P(v_x, v_y, v_z) = n \left(\frac{m}{2\pi\kappa_B T} \right)^{3/2} e^{-m \left[(v_x - u)^2 + v_y^2 + v_z^2 \right] / 2\kappa_B T}$$

From $f = f_0 - \tau (\mathbf{v} \cdot \nabla_r f_0)$ (\mathbf{F} is zero because this is no external fields here and gravity is neglected), the distribution function is

$$f = f_0 - \tau (v_x \hat{i} + v_y \hat{j} + v_z \hat{k}) \cdot \left(\frac{\partial f_0}{\partial x} \hat{i} + \frac{\partial f_0}{\partial y} \hat{j} + \frac{\partial f_0}{\partial z} \hat{k} \right) = f_0 - \tau v_y \frac{\partial f_0}{\partial y} = f_0 - \tau v_y \frac{\partial f_0}{\partial u} \frac{\partial u}{\partial y}$$

Compared with the law $\tau_{xy} = \mu \frac{\partial u}{\partial y}$ (x is the force direction, y is the normal direction), we

have

$$\mu = n\tau k_B T$$

Following a similar procedure, we can also calculate the energy flux due to molecular heat conduction and obtain the thermal conductivity for a gas as

$$k = \frac{5}{2} \left(\frac{k_B}{m} \right) n\tau k_B T = \frac{5}{2} \left(\frac{k_B}{m} \right) \mu$$

Note: The thermal conductivity and dynamic viscosity have similar form here because they are based on the same mechanism.

In the following lecture, we will discuss the Ohm's Law and Wiedmann-Franz's Law. The force acting on the electron by the external field is

$$\mathbf{F} = -e\boldsymbol{\mathcal{E}},$$

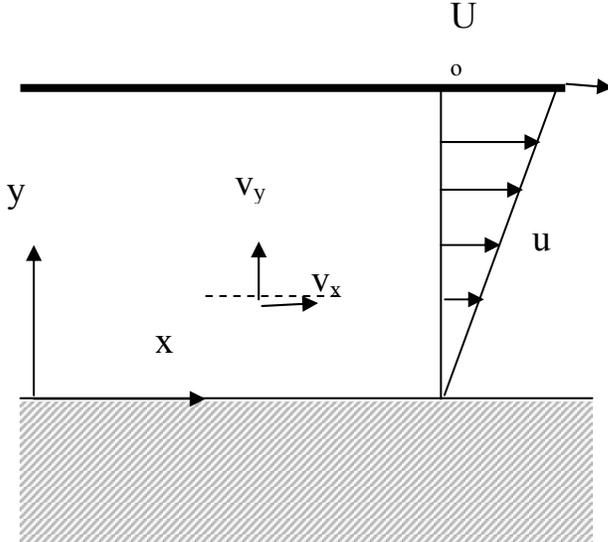
where e is the unit charge, and the charge of an electron is $(-e)$, $\boldsymbol{\mathcal{E}}$ is the electric field. And f_0 obeys the Fermi-Dirac distribution

$$f_0(E, E_f, T) = \frac{1}{\exp\left(\frac{E - E_f}{\kappa_B T}\right) + 1}.$$

We will deal with μ , T and F simultaneously in this situation.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 19

In last lecture, we talked about the Newton's shear stress law.



Assuming that the number density of particles is n , the number density of particles having velocity \mathbf{v} is

$$f_0(v_x, v_y, v_z) = nP(v_x, v_y, v_z) = n \left(\frac{m}{2\pi\kappa_B T} \right)^{3/2} e^{-m \left[(v_x - u)^2 + v_y^2 + v_z^2 \right] / 2\kappa_B T}$$

From $f = f_0 - \tau \left(\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_0 \right)$ (\mathbf{F} is zero here because this is no external fields here and gravity is neglected), the distribution function is

$$f = f_0 - \tau \left(v_x \hat{i} + v_y \hat{j} + v_z \hat{k} \right) \cdot \left(\frac{\partial f_0}{\partial x} \hat{i} + \frac{\partial f_0}{\partial y} \hat{j} + \frac{\partial f_0}{\partial z} \hat{k} \right) = f_0 - \tau v_y \frac{\partial f_0}{\partial y} = f_0 - \tau v_y \frac{\partial f_0}{\partial u} \frac{\partial u}{\partial y}$$

Note: We can also prove $u = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v_x f dv_x dv_y dv_z$.

The shear stress along the x-direction, on a plane perpendicular y-axis can be calculated by considering the momentum exchange across the plane,

$$\begin{aligned} \tau_{xy} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v_y [m v_x] f dv_x dv_y dv_z \\ &= -\frac{\partial u}{\partial y} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tau v_y^2 m v_x \frac{\partial f_0}{\partial u} dv_x dv_y dv_z = \mu \frac{\partial u}{\partial y} \end{aligned}$$

The dynamic viscosity can be

$$\mu = -\int \int \int \tau v_y^2 (m v_x) \frac{\partial f_0}{\partial u} dv_x dv_y dv_z$$

$$= m^2 \tau n \left(\frac{m}{2\pi\kappa_B T} \right)^{3/2} \int_{-\infty}^{\infty} e^{-mv_z^2/(2\kappa_B T)} dv_z \int_{-\infty}^{\infty} v_y^2 e^{-mv_y^2/(2\kappa_B T)} dv_y \int_{-\infty}^{\infty} \frac{v_x^2}{\kappa_B T} e^{-mv_x^2/(2\kappa_B T)} dv_x' .$$

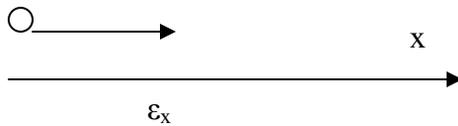
Note: Two integrations are used to calculate the above equation:

$$(1) \left(\int_{-\infty}^{+\infty} e^{-x^2} dx \right)^2 = \left(\int_{-\infty}^{+\infty} e^{-x^2} dx \right) \left(\int_{-\infty}^{+\infty} e^{-y^2} dy \right) = \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy = \int_0^{+\infty} e^{-r^2} 2\pi r dr \text{ yields}$$

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi} .$$

$$(2) f(a) = \int_{-\infty}^{+\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \text{ leads to } f'(a) = -\int_{-\infty}^{+\infty} x^2 e^{-ax^2} dx = \left(\sqrt{\frac{\pi}{a}} \right)' = -\frac{\sqrt{\pi}}{2} a^{-3/2} .$$

Ohm's Law

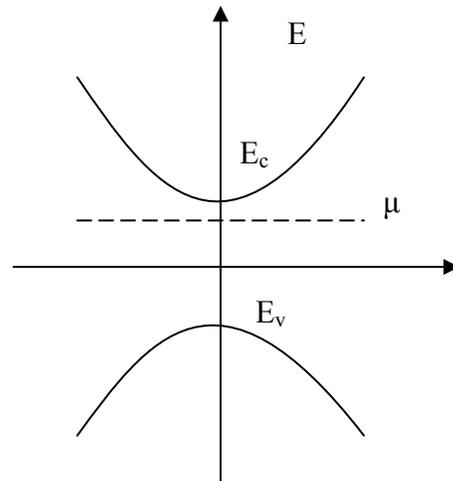
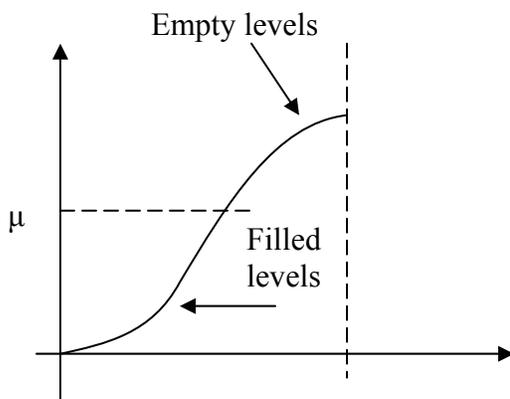


Now f_0 obeys the Fermi-Dirac distribution

$$f_0(E, E_f, T) = \frac{1}{\exp\left(\frac{E - \mu}{\kappa_B T}\right) + 1} ,$$

and $f = f_0 - \tau \left(\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_0 \right)$ becomes

$$f = f_0 - \tau \left(v_x \frac{df_0}{dx} + \frac{q\epsilon_x}{m} \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial v_x} \right) .$$



For metals (left figure), the electrons fill part of the band, while the Fermi level normally lies within the band gap for semiconductors. In the latter situation, the parabolic approximation is used

$$E = E_c + \frac{1}{2} \frac{\hbar^2 k^2}{m^*} = E_c + \frac{1}{2} m^* (v_x^2 + v_y^2 + v_z^2),$$

thus

$$f = f_0 - \tau \left(v_x \frac{\partial f_0}{\partial x} + \frac{q \varepsilon_x}{m} \frac{\partial f_0}{\partial E} \frac{\partial E}{\partial v_x} \right) = f_0 - \tau v_x \left(\frac{\partial f_0}{\partial x} + q \varepsilon_x \frac{\partial f_0}{\partial E} \right).$$

The current density (A/m²) is then

$$J_q = \frac{1}{V} \sum_{k_x=-\infty}^{\infty} \sum_{k_y=-\infty}^{\infty} \sum_{k_z=-\infty}^{\infty} q v_x f = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} q v_x f dk_x dk_y dk_z$$

Changing to spherical coordinate system ($dk_x dk_y dk_z = k^2 \sin \theta d\theta d\varphi dk$) and rewrite the above equation with $D(E)$ yields

$$\begin{aligned} J_q &= \int_0^{2\pi} \int_0^{\pi} \int_0^{+\infty} q v_x f \frac{D(E)}{4\pi} \sin \theta d\theta d\varphi dE \\ &= -\frac{q}{3} \int_0^{\infty} \tau v^2 D(E) \left(\frac{\partial f_0}{\partial x} + q \varepsilon_x \frac{\partial f_0}{\partial E} \right) dE \end{aligned}$$

Now we let $E_c=0$ be the reference energy level. And the distribution becomes

$$f_0(E, E_f, T) = \frac{1}{\exp\left(\frac{E_k - (\mu - E_c)}{\kappa_B T}\right) + 1},$$

where E_k denotes the kinetic energy.

Noticing $\frac{\partial f_0}{\partial T} = -\frac{E_k - \mu}{T} \frac{\partial f_0}{\partial E_k}$, $\frac{\partial f_0}{\partial \mu} = -\frac{\partial f_0}{\partial E_k}$, we have

$$\begin{aligned} \frac{\partial f_0}{\partial x} &= \frac{\partial f_0}{\partial \mu} \frac{\partial \mu}{\partial x} + \frac{\partial f_0}{\partial T} \frac{\partial T}{\partial x} \\ &= -\frac{\partial f_0}{\partial E_k} \frac{\partial \mu}{\partial x} - \frac{E_k - \mu}{T} \frac{\partial f_0}{\partial E_k} \frac{\partial T}{\partial x} \end{aligned}$$

Note: Defining $y = \frac{E_k - \mu}{\kappa_B T}$, $f_0(E, E_f, T) = \frac{1}{\exp(y) + 1}$, we can observe

$$\frac{\partial f_0}{\partial T} = -\frac{E_k - \mu}{T} \frac{\partial f_0}{\partial E_k}$$

from

$$\frac{\partial f_0}{\partial T} = -\frac{\partial f_0}{\partial y} \frac{E_k - \mu}{\kappa_B T^2}, \quad \frac{\partial f_0}{\partial E_k} = \frac{\partial f_0}{\partial y} \frac{1}{\kappa_B T}.$$

The flux is thus

$$J_q = -\frac{q}{3} \int_0^\infty \tau v^2 \left(-\frac{\partial \mu}{\partial x} - \frac{E_k - \mu}{T} \frac{\partial T}{\partial x} + q \varepsilon_x \right) \frac{\partial f_0}{\partial E_k} D(E_k) dE_k$$

$$= \left\{ \frac{1}{3} \int_0^\infty q \tau v^2 \frac{\partial f_0}{\partial E_k} D(E_k) dE_k \right\} \left(\frac{\partial \mu}{\partial x} - q \varepsilon_x \right) + \left\{ \frac{1}{3} \int_0^\infty q \tau v^2 \left(\frac{E_k - \mu}{T} \right) \frac{\partial f_0}{\partial E_k} D(E_k) dE_k \right\} \frac{\partial T}{\partial x}$$

Discussion

To simplify, let $L_{11} = \frac{1}{3} \int_0^\infty q \tau v^2 \frac{\partial f_0}{\partial E_k} D(E_k) dE_k$.

1) $\partial T / \partial x = 0$ (isothermal), we have

$$J_x = L_{11} \left(\frac{\partial \mu}{\partial x} - q \varepsilon_x \right) = L_{11} q \frac{\partial \phi}{\partial x},$$

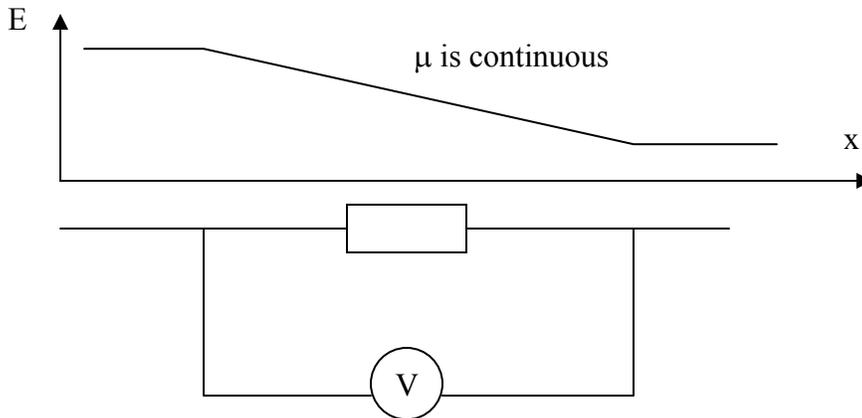
where $\phi = \frac{\mu}{q} + v$ is the electrochemical potential.

Note: $\frac{\partial \mu}{\partial x} - q \varepsilon_x = \frac{\partial \mu}{\partial x} + q \frac{dV}{dx} = \frac{\partial}{\partial x} \left(\frac{\mu}{q} + v \right)$.

The flux is often written as

$$J_x = qn\mu_e \varepsilon_x + Dq \frac{\partial n}{\partial x}$$

where the first term corresponds to drift, the second term denotes diffusion; μ_e is called the **mobility** [$\text{m}^2\text{V}^{-1}\text{s}^{-1}$] and D is the **diffusivity** [m^2s^{-1}], n is carrier concentration. In the following figures, we actually measure both terms in the above equation.



Note: Recall we use $n = \int fD(E)dE$ to determine μ in Lecture 11. Normally n is related to μ . In semiconductors, when we increase doping, n will increase while μ will drop due to scattering. In metals, the large n is almost constant and $J_e = qn\mu_e \varepsilon_x = \sigma \varepsilon_x$.

In non-degenerate semiconductors, the Fermi-Dirac distribution can be approximated by the Boltzmann distribution and $D = k_B T \mu / q$, called Einstein relationship. This name comes because it is derived following Einstein's work on Brownian motions.

2) $\partial T / \partial x \neq 0$

Now the equation is written as

$$J_x = L_{11} q \left(\frac{d\Phi}{dx} \right) + L_{21} \left(\frac{dT}{dx} \right).$$

For open circuits (zero flux), we have

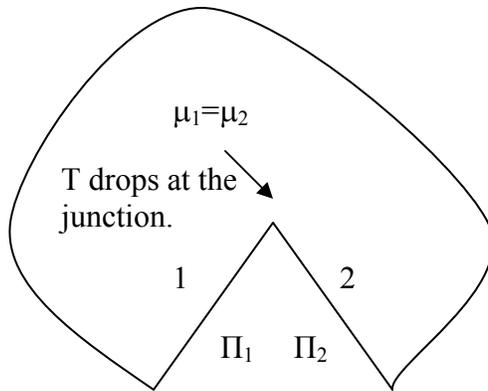
$$0 = L_{11} q \left(\frac{d\Phi}{dx} \right) + L_{21} \left(\frac{dT}{dx} \right),$$

thus

$$\frac{d\Phi}{dx} = - \frac{L_{12}}{L_{11}} \frac{dT}{dx} = -S \frac{dT}{dx},$$

where $S = \frac{L_{12}}{L_{11}}$ (VK⁻¹) is called **Seebeck coefficient**. The Seebeck voltage is the steady-

state voltage accumulated under the open circuit condition. If the conductor is a uniform material such that S is a constant, the voltage difference does not depend on the temperature profile. This is the principle behind the thermocouple for temperature measurements. A thermocouple (the following figure) employs two conductors for the easy of measuring the voltage difference.



For the heat flux, we have

$$dE = dq + \mu dn,$$

$$J_q = \sum \sum \sum 2(E - \mu) v_x f = L_{21} \frac{d\phi}{dx} + L_{22} \frac{dT}{dx}.$$

The heat flux is also written as $J_q = \Pi J_x$, in which the Peltier factor $\Pi = ST$.

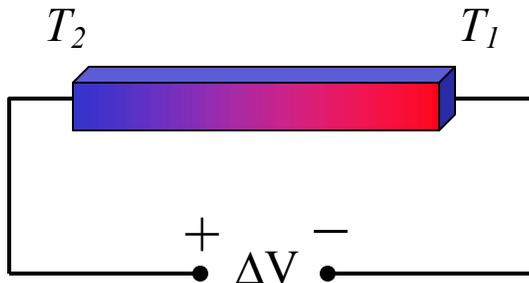
2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 20

Guest lecture by Prof. Mildred S. Dresselhaus

1. Outline

- Overview of low dimensional thermoelectricity
- New physics to yield enhanced performance in 1D nanowires
- Quantum dot superlattice nanowires – model calculations
- Newly emerging research directions
- New methods for synthesis and assembly of nanowires
 - Self assembled composite nanostructures
 - New 3D crystalline materials with quantum dots
 - New thermoelectric tools

2. Introduction to Thermoelectricity

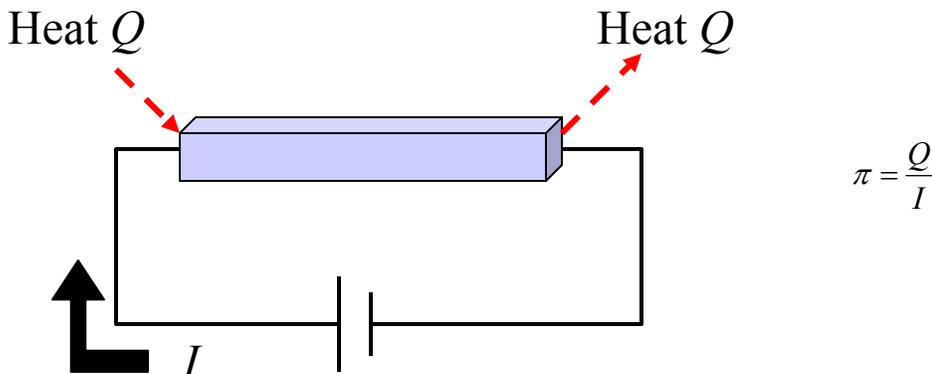


Physically, when one side of the conductor (or semiconductor) is hot, electrons have higher thermal energy and will diffuse to the cold side. The higher charge concentration in the cold side builds an internal electric field that resists the diffusion. The Seebeck voltage is the steady-state voltage accumulated under the open circuit condition. We can express this as

$$S = -\frac{\Delta V}{T_2 - T_1} \text{ (V/K)},$$

where $S > 0$ for p-type semiconductors, and $S < 0$ for n-type materials.

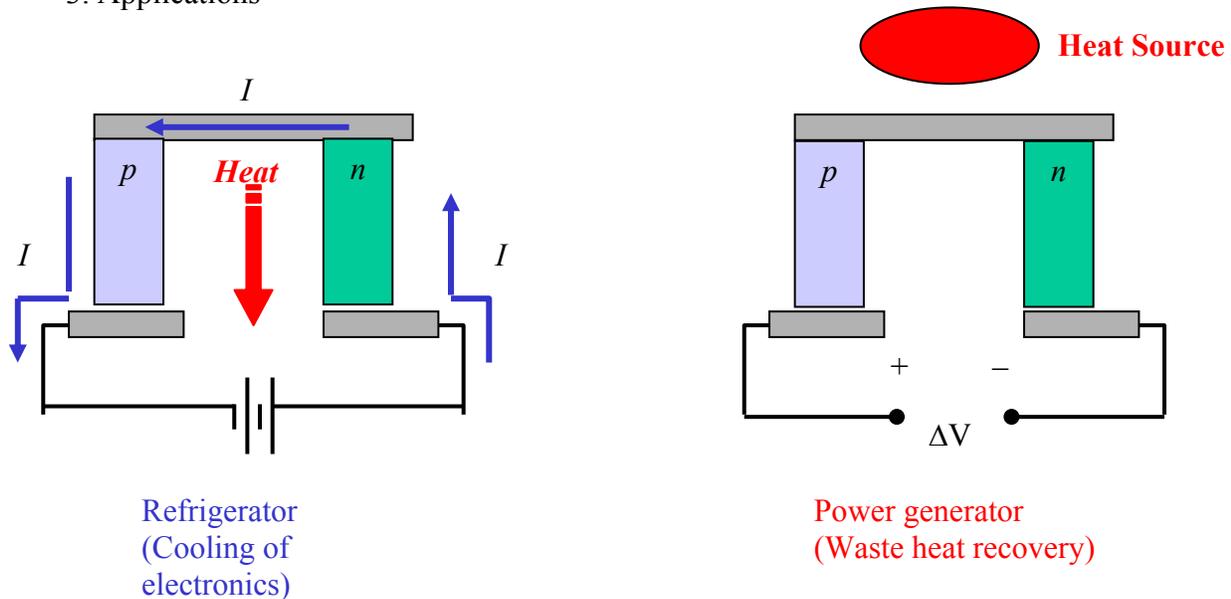
In the following figure, we demonstrate the idea of Peltier Effects.



$$\pi = \frac{Q}{I}$$

In this case, heat is carried by carriers (absorbed at the left and rejected at the right end). The Peltier coefficient is defined as $\Pi = Q/I$.

3. Applications



In the above figure, we show two major applications of thermoelectricity: refrigerator and power generator. The efficiency of the refrigerator can be evaluated by $ZT = \frac{S^2 \sigma}{k} T$, in

which S is seebeck coefficient, σ is electrical conductivity, and k denotes thermal conductivity. Compared with mechanical refrigeration, thermoelectric cooling offers the following advantages:

- No moving parts
- Environmentally friendly
- No loss of efficiency with size reduction
- Can be integrated with electronic circuits (e.g. CPU)
- Localized cooling with rapid response

However, the current ZT value (~ 1) for TE cooling still lags behind that of mechanical refrigeration ($ZT \sim 3$).

4. Thermoelectric Properties of Conventional Materials

In the following figure, we present the variations of three parameters against the carrier concentration. The ZT value reaches the maximum in the middle of the figure. However, we should not use semimetals because they have both holes and electrons as carriers, which have different signs of S and the effects will cancel out in the cooling process. Therefore, we should focus on the edge of semiconductors.

To increase ZT , we want $S \uparrow, \sigma \uparrow, k \downarrow$. There is conflict in satisfying all these requirements. In the figure, we can find $S \uparrow, \sigma \downarrow$ and $\sigma \uparrow, k \uparrow$. Basically we need to balance these factors. The best alloy $\text{Bi}_{0.5}\text{Sb}_{1.5}\text{Te}_3$ has $ZT \sim 1$ at 300 K.

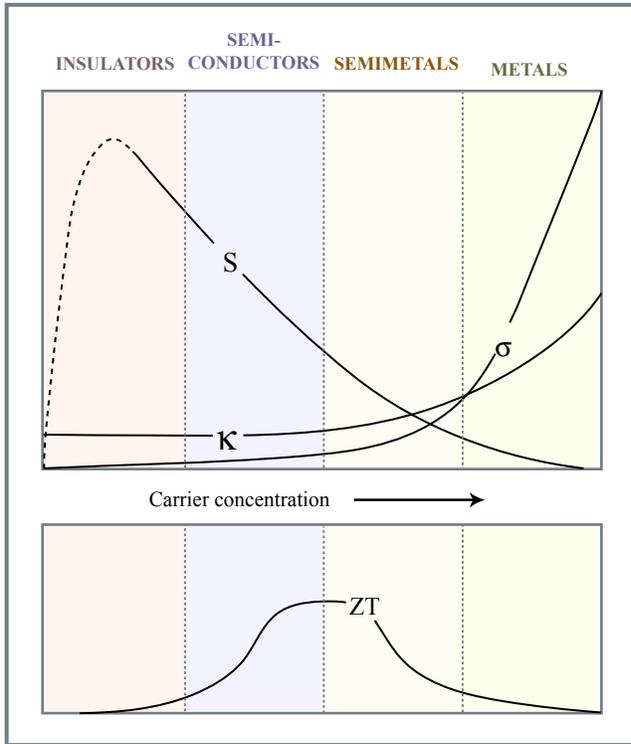


Image by MIT OpenCourseWare.

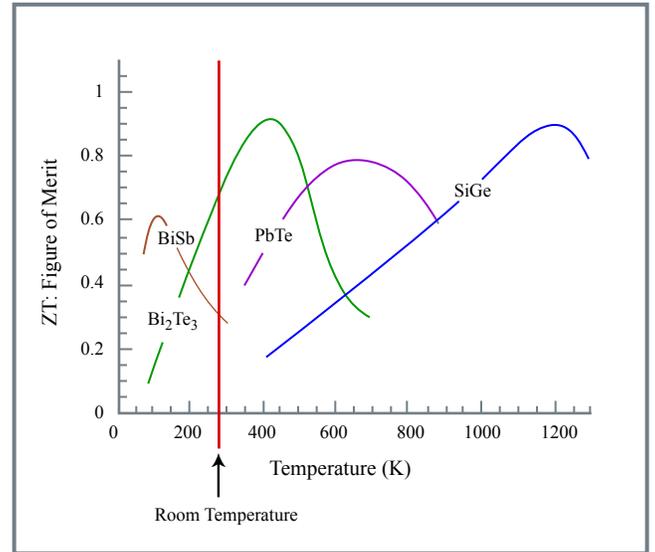


Image by MIT OpenCourseWare.

In the current investigations, people have tried different compositions to improve ZT . At different temperature ranges, in the right figure we have different best materials. The ZT values are still less than 1.0 after 40 years.

5. Motivation for Nanotech Thermoelectricity

To further improve the efficiency, low dimensional materials give additional control in

- Enhanced density of states due to quantum confinement effects (we can increase S without reducing σ)
- Boundary scattering at interfaces reduces k more than σ

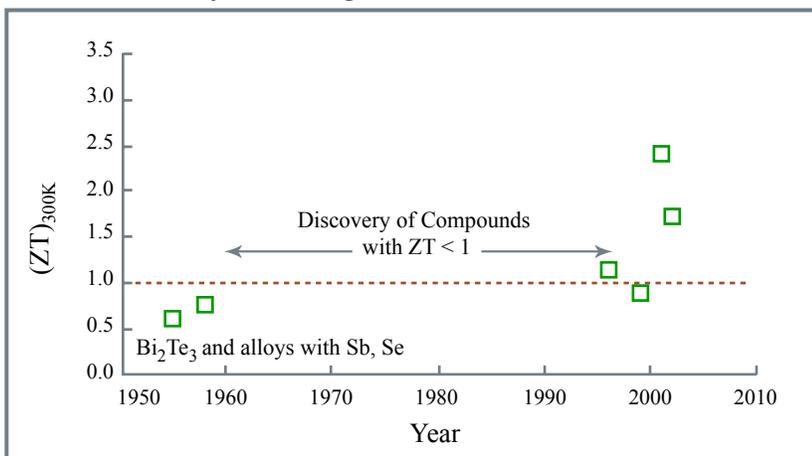
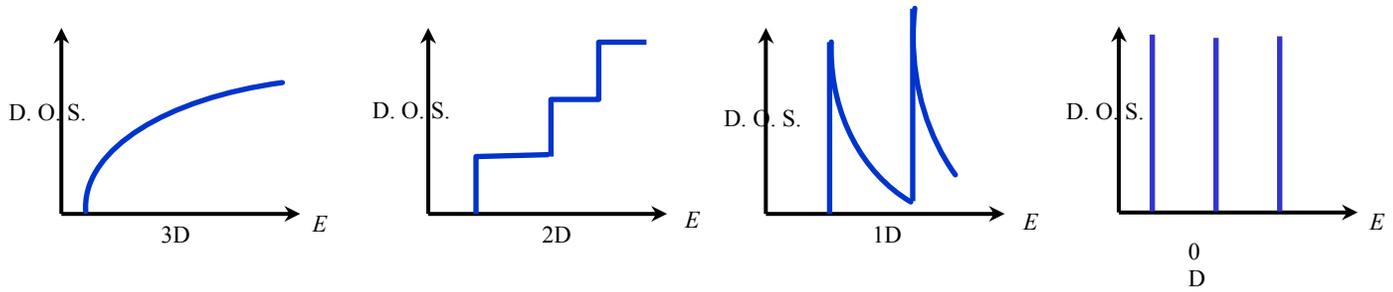


Image by MIT OpenCourseWare.

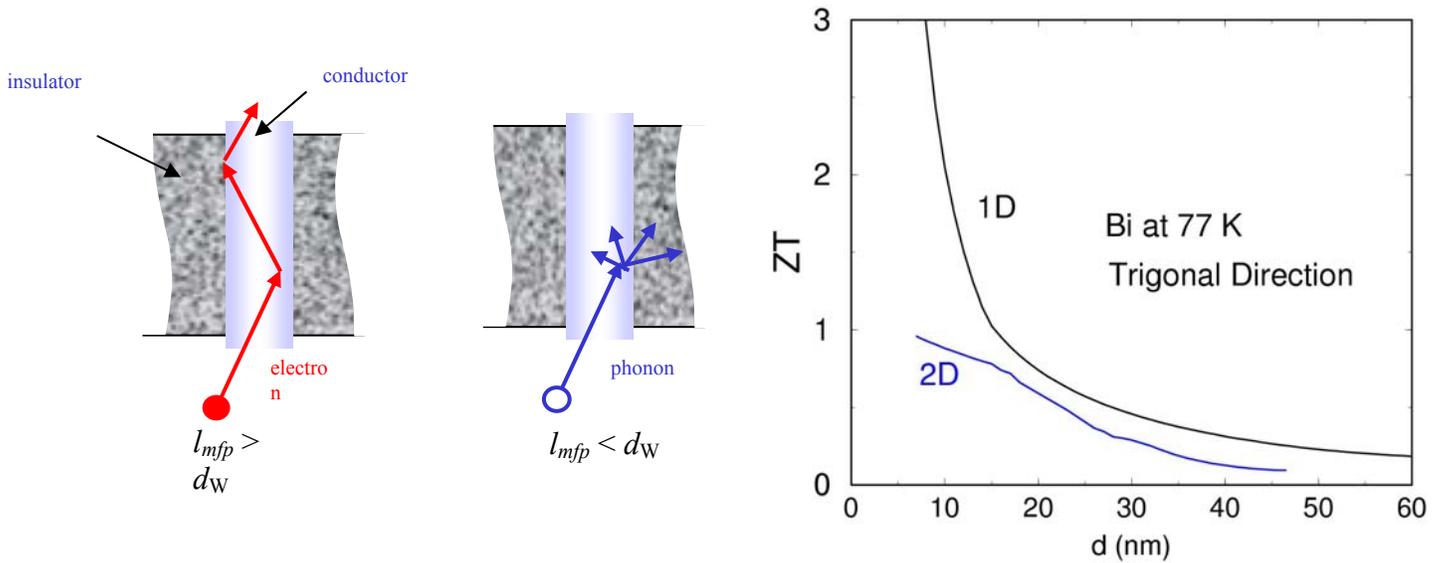
The above figure summarizes the recent advances in nanostructured thermoelectric materials, which led to a sudden increase in $(ZT)_{300K} > 1$. Higher ZT reported experimentally at higher T.

6. New Directions for Low Dimensional or Nanotech Thermoelectricity

1) Electronic properties may be dramatically modified due to the electron confinement in nanostructures, which exhibit low-dimensional behaviors. Recall the following cases given in homework.



2) Thermal conductivity can be significantly reduced by the scattering of phonons at the interfaces. In the following case, the electrical conductivity is not strongly affected.



In the above right figure, the 1D structure shows promise of increasing ZT to 3 at a 10 nm diameter. In the next two figures, thermoelectric cooling using a Bi_2Te_3/Sb_2Te_3 superlattice structure is reported. Enhanced cooling is accomplished by increased scattering of phonons at interfaces, thus lowering the lattice thermal conductivity.

Figure removed for copyright reasons.

Venkatasubramanian *et al.*, Nature **413**, 589 (2001)

Figure removed for copyright reasons.

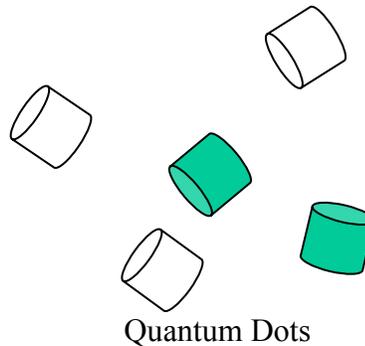
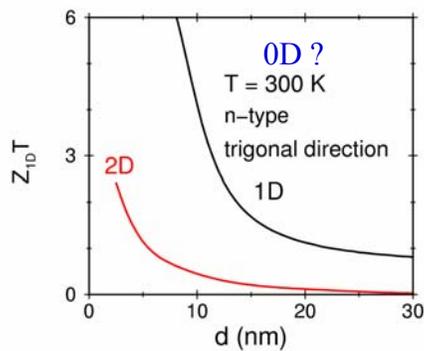
Harman *et al.* J. Electron. mater. Lett. **29**, L1 (2000)

In the above work, enhancement of ZT at 300 K in a quantum dot superlattice is more than a factor of two relative to best available bulk PbTe because

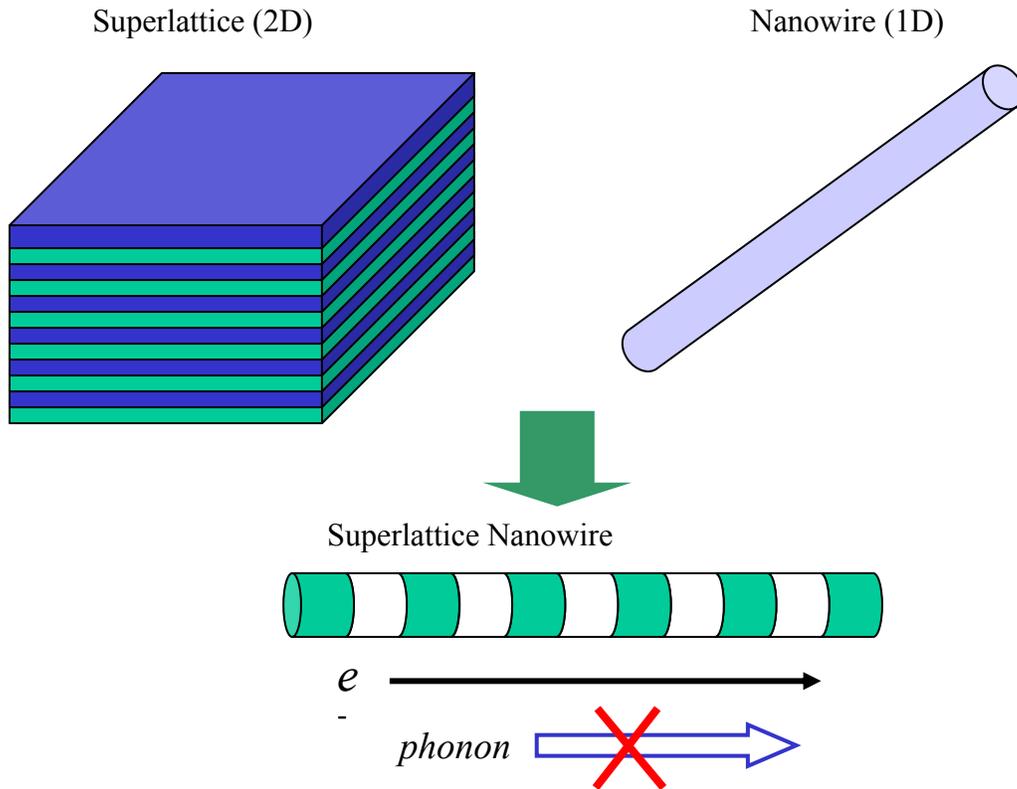
- Favorable carrier scattering mechanism due to PbSeTe quantum dots
- Lower lattice thermal conductivity in alloy

And ZT = 3 at 600 K reported at MRS Boston in 2003.

7. Quantum dot superlattice nanowires – model calculations



We have discussed the seebeck coefficient enhancement in 2D and 1D structure. Reducing the dimension to 0D may further increase S . However, difficulty exists in making connections to different quantum dots. This problem is solved by combining the idea of superlattice and nanowires. The new structure (superlattice nanowire) will restrict heat conduction but allow electrons to pass.



In the following figures, three superlattice nanowires are presented.

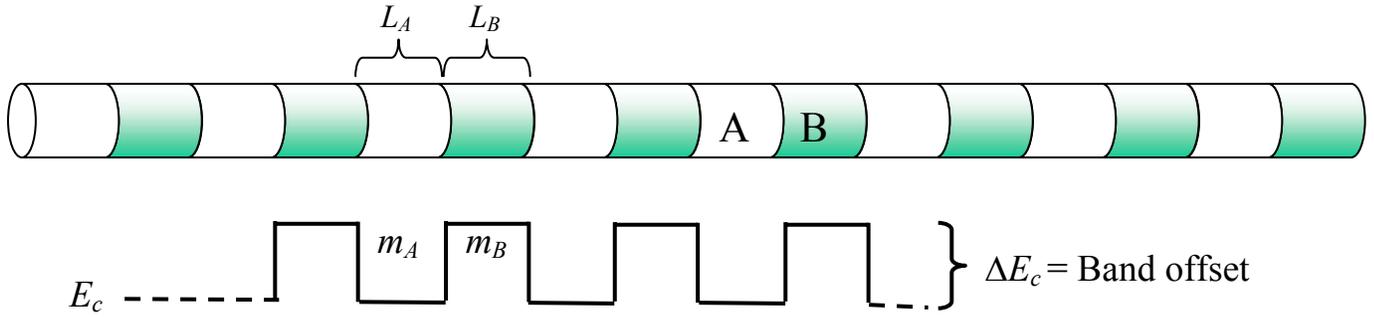
Photo removed for copyright reasons. See Figure 3 in Piraux et al. "Giant magnetoresistance in magnetic multilayered nanowires." *Appl. Phys. Lett.* 65, no. 2484 (1994).

Photo removed for copyright reasons. See Fig. 2b in Wu, Fan, and Yang. "Block-by-block growth of single-crystalline Si/SiGe superlattice nanowires." *Nano Lett.*, 2 (2002): 83.

Photo removed for copyright reasons. See Figure 1d in Bjork, M.T., et al. "One-dimensional Steeplechase for Electrons Realized." *Nano Lett.* 2: 2 (2002): 87-89.

The theoretical modeling of this structure includes the following key parameters

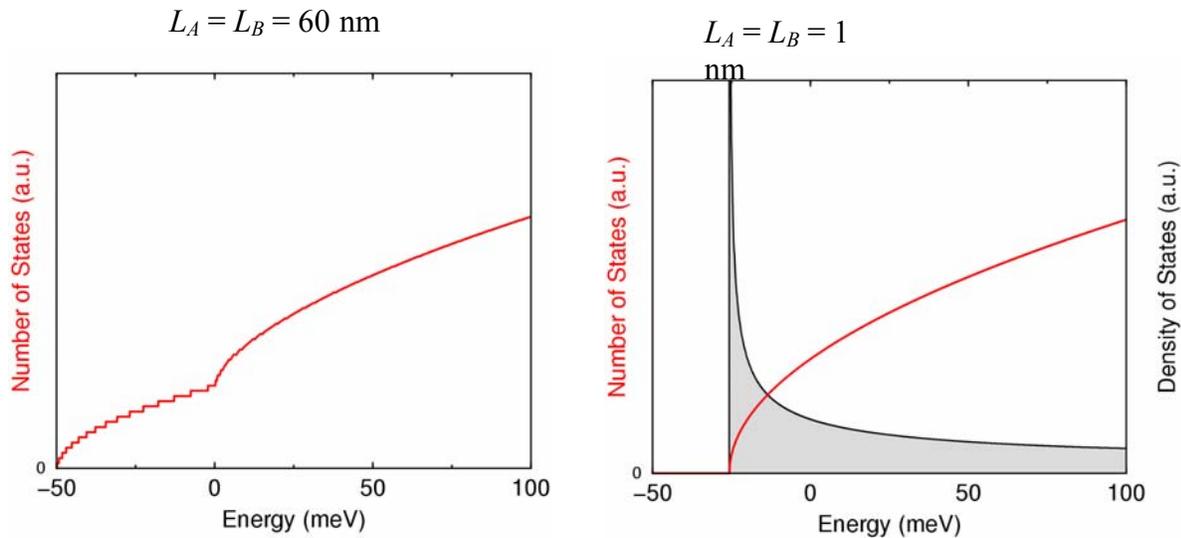
- Wire diameter: d_W
- Segment lengths: L_A and L_B (assumed to be equal)
- Subband potential barrier: ΔE_c
- Effective masses: m_A and m_B
- Phonon mean free path: L



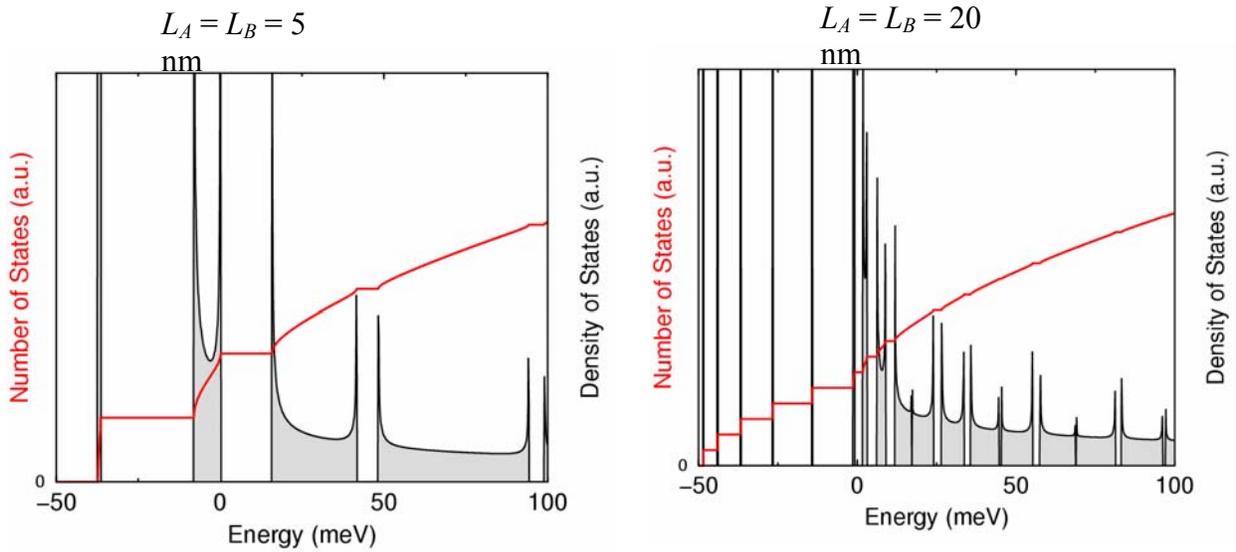
Y-M. Lin and M.S. Dresselhaus., Phys. Rev. B **68**, 0753045 (2003)

The utilized approaches are

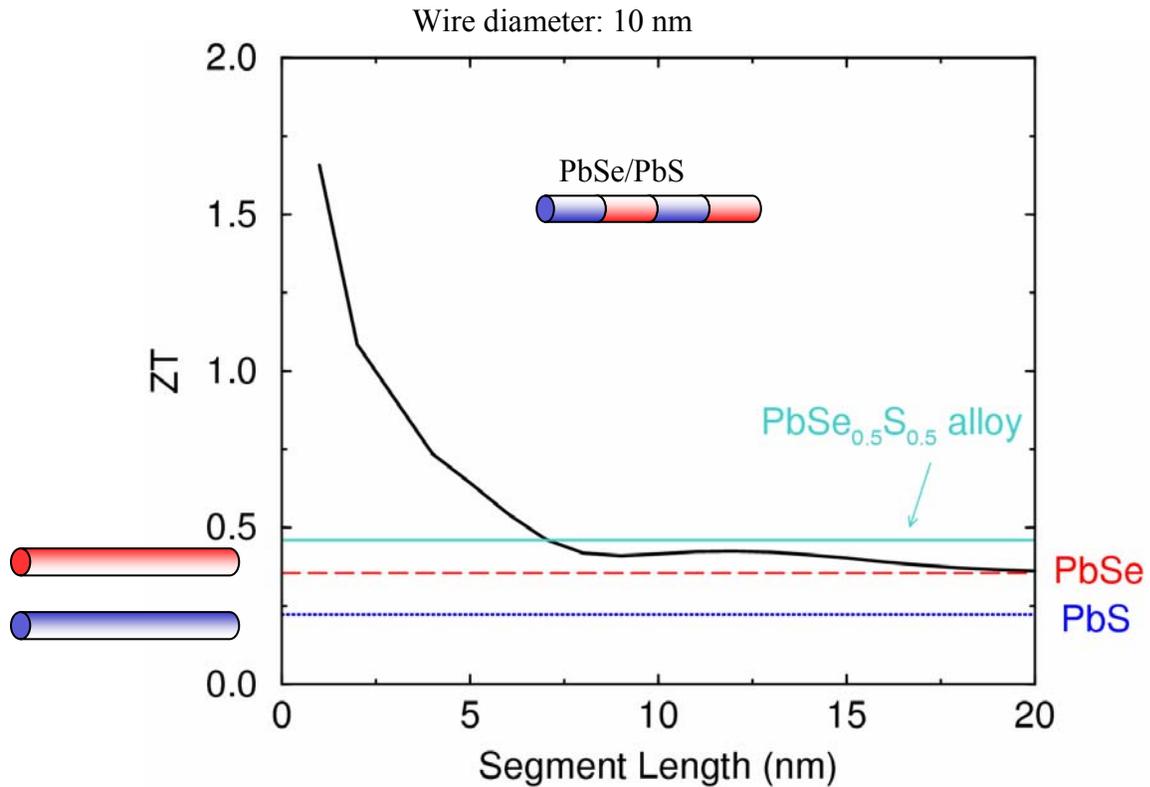
- Determination of the (sub)band structure
- Derivation of the dispersion relation $E(k)$ along the wire axis
- Calculation of thermoelectric properties based on Boltzmann transport equations



In the above figures, the DOS for classical limit and alloy limit are shown. In real applications, we want more sharp pulses (~ 10) but not too many. In the following figures, the left one is suitable, while the right one already shows the trend to be a 3D structure.



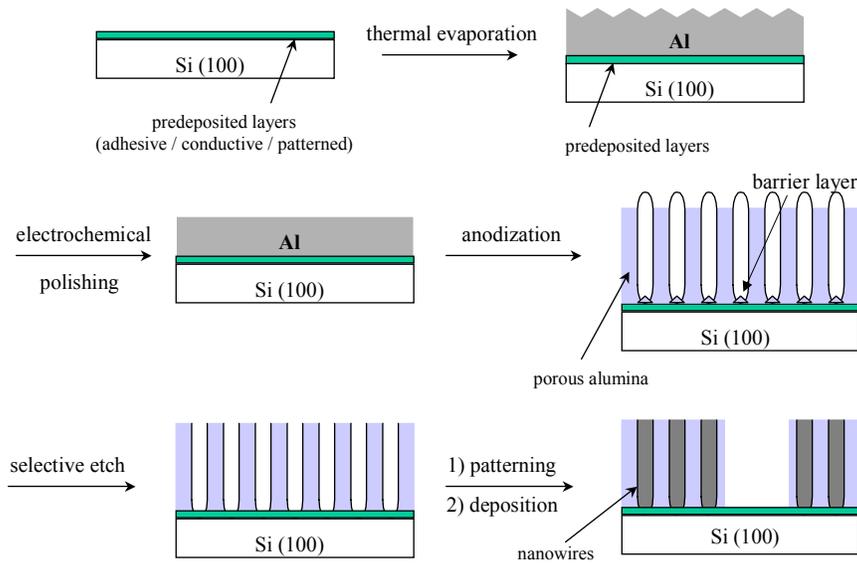
The following presents ZT for [001] n-type PbSe/PbS SL nanowires as a function of segment length at 77 K. Greater enhancement is predicted for SL nanowires with diameters of 5nm.



8. Newly emerging research directions

- 1) New methods for synthesis and assembly of nanowires

The following figure demonstrate the idea of making standing nanowires.

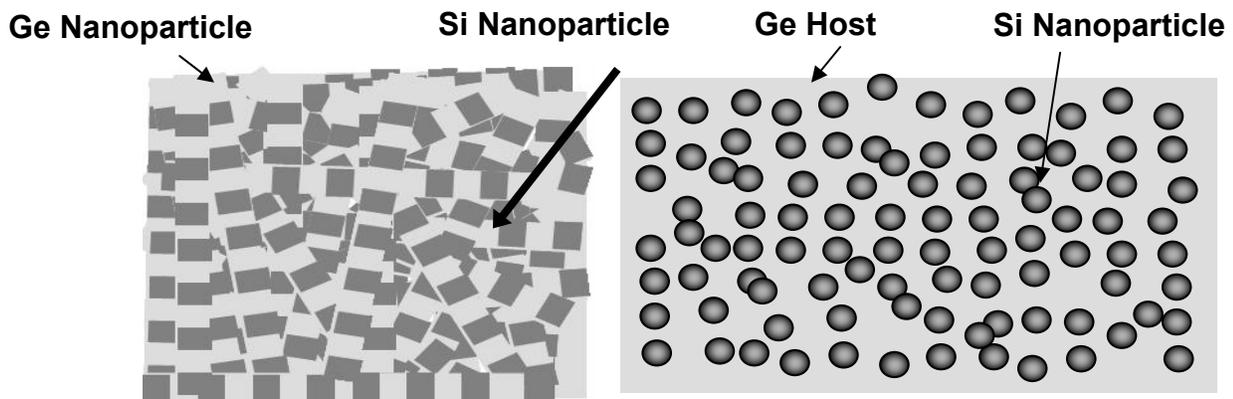


Two pictures of nanowires are shown below.

Photo removed for copyright reasons.

Rabin *et al.*, *Adv. Funct. Mater.* (2003)

2) Nanoparticle Composite Synthesis



Gang Chen, MIT

In the above figures, the idea of using nanoparticle composites is demonstrated. Expectations include:

- Reduced thermal conductivity.
- Electrical conductivity comparable to or better than bulk.
- Increased thermoelectric figure of merit.
- Cheap, self-assembly method

3) New Nanoscale Tools for Thermoelectric Measurements

Scanning thermoelectric microscopy (SThEM) allows measurements of the spatial profiles of the thermoelectric voltage, carrier concentration and electron energy band of a p-n junction to 2 nm resolution. The following work has general implications on electronics and opto-electronics of nano-systems.

Diagram and graphs removed for copyright reasons.

SThEM measurement set-up

Thermoelectric voltage profile across p-n junction with 2nm resolution

Carrier profiles and electronic band energies across p-n junction

H.-K. Lyeo et al. , Science 303, 816 (2004)

9. Conclusions

1) Model systems show that:

- ZT for 0D nanowire superlattice
- > ZT for 1D quantum wires
- > ZT for 2D quantum wells
- > ZT for bulk for same material

2) New research directions now being pursued:

- Self assembled bulk composites of nanostructures
- New 3D crystalline materials with quantum dots
- New thermoelectric tools at the nanoscale

3) Objective

- To have compact technology for cooling, especially for electronics and opto-electronics
- To have efficient method for converting thermal to electrical energy, including waste heat recovery

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 21

Last time we talked about the current density as

$$J_e = L_{11} 'q \left(\frac{d\Phi}{dx} \right) + L_{12} \left(\frac{dT}{dx} \right),$$

For electrons, $q = -e$ and $\Phi = \varphi_e - \frac{\mu}{e}$. Here φ_e is electrostatic potential, which is related to the electrical field. Chemical potential μ is related to diffusion. Their combination Φ is electrochemical potential, indicating the total driving force of charges. The current density can be rewritten as

$$J_e = L_{11} \left(-\frac{d\Phi}{dx} \right) + L_{12} \left(\frac{dT}{dx} \right).$$

Note: The second term $L_{12} \left(\frac{dT}{dx} \right) = L_{12} T \left(\frac{1}{T} \frac{dT}{dx} \right)$ is similar to $\Delta S = \frac{dQ}{T}$ and may be compared with entropy flux.

The heat transferred is

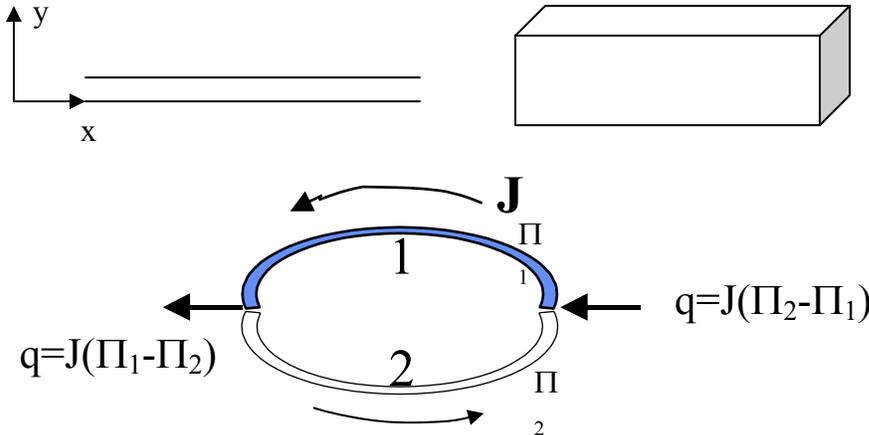
$$J_q = \frac{2}{V} \sum_{k_x} \sum_{k_y} \sum_{k_z} v_x (E - \mu) f = L_{21} \left(-\frac{d\Phi}{dx} \right) + L_{22} \frac{dT}{dx}$$

For open circuits, $J_e=0$. We obtain

$$-S = \frac{d\Phi / dx}{dT / dx} = \frac{V}{T_h - T_c} = \frac{L_{12}}{L_{11}},$$

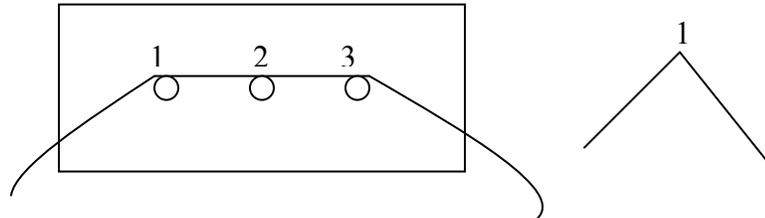
where S is called the seebeck coefficient.

Note: (1) S is dependent on the density of states. Therefore, it can be enhanced by using nanostructures such as thin films or nanowires. This effect is also the principle of thermal couples. (2) In the summation, we cannot use integral for quantized directions.



(2) In $S = -\frac{d\Phi/dx}{dT/dx}$, T should be the electron temperature T_e . In equilibrium cases, T_e is close to phonon T_p and we can use this effect to measure T_p . However, for extreme cases such as laser ablation, the two temperatures are not in equilibrium. Cautions should be taken.

(3) For an on-chip thermocouple, the measured temperature does not correspond to the junction point, but closer to the average temperature from 1 to 3. This is different from the normal thermocouples.



When $dT/dx=0$, we have

$$J_q = L_{21} \left(-\frac{d\Phi}{dx} \right) = \frac{L_{21}}{L_{11}} J_e = \Pi J_e,$$

where the Peltier coefficient $\Pi = TS$, $L_{21}=TL_{12}$. Note one thermoelectric coefficient (S here) can be used to express all other coefficients. This is a requirement of the “time reversal invariance” of the mechanical equations of motion, i.e., the particles retrace their former paths if all velocities are reversed. Based on this principle, Onsager (1931) derived the famous **Onsager reciprocity relations**. The flux of any extensive variable, J_k , of a system (such as energy flux, particle flux) or at a local point of a system can be expressed as a linear combination of all the generalized driven forces F_j ,

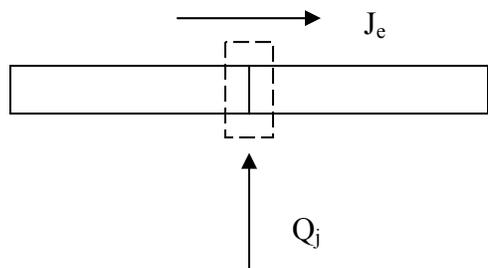
$$J_k = \sum_j L_{jk} F_j.$$

Onsager got a Nobel Prize for his work.

Now apply the above equation to the small control volume shown in the following figure. The energy conservation yields

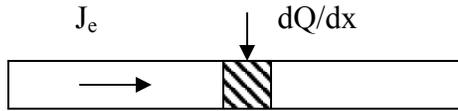
$$Q_j = J_{q2} - J_{q1} = (\Pi_2 - \Pi_1) J_e,$$

which is positive (absorbing heat from the ambient) for $\Pi_2 > \Pi_1$.



Another phenomenon is the Thomson effect. Along the following bar, we have

$$T \frac{dS}{dT} = \frac{1}{J_e} \frac{dq/dx}{dT/dx} = \beta.$$



The overall energy equation is

$$q = \frac{dJ_q}{dx} + J_e \frac{d\Phi}{dx} = \sigma J_e^2 + k \frac{dT}{dx} + \text{Thomson term}.$$

And thermal conductivity is

$$k_e = L_{12}L_{21} / L_{11} - L_{22}.$$

The Wiedemann-Franz law states

$$\frac{k_e}{\sigma T} = \text{const} = L = 2.45E-8 \text{ W}\Omega / \text{K}^2,$$

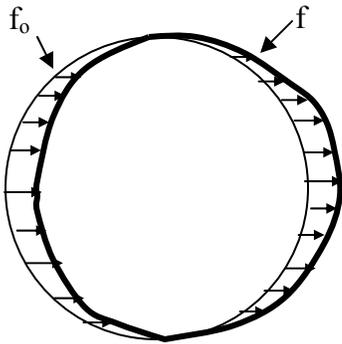
where the constant L is Lorenz number.

Now consider the Boltzmann equation under the relaxation time approximation

$$\frac{\partial g}{\partial t} + \frac{\partial f_0}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 + \mathbf{v} \cdot \nabla_{\mathbf{r}} g + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_0 + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} g = -\frac{g}{\tau}, \quad g = f - f_0.$$

By neglecting some terms, we have obtained

$$g = -\tau \left(\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_0 \right).$$



(1) Assume that f only has small deviation from f_0 , indicating local equilibrium. Now we have $g = f - f_0 \ll f_0$. This yields

$$\tau \mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 \ll f_0$$

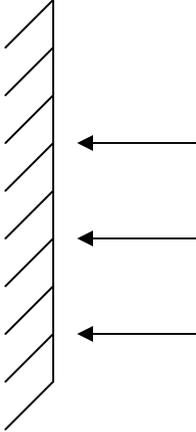
$$\text{or } \tau \mathbf{v} \frac{df_0}{dT} \frac{dT}{dx} \ll f_0 \Rightarrow \frac{\Lambda}{T} \frac{dT}{dx} \ll 1 \Rightarrow \Lambda / L \ll 1, \quad \frac{dT}{dx} \text{ should not be too large.}$$

(2) Within $\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0, \mathbf{v} \cdot \nabla_{\mathbf{r}} g$, we ignore the latter one. This implies

$$\nabla_{\mathbf{r}} f_0 \gg \nabla_{\mathbf{r}} g$$

(3) We ignore $\frac{\partial f_o}{\partial t}$ on the LHS compared with $\frac{g}{\tau}$ on the RHS. This indicates

$$\frac{\partial f_o}{\partial t} \ll \frac{g}{\tau} \Rightarrow t \gg \tau.$$



Now consider the above transient process in which an infinite wall is heated suddenly.

The theoretical solution of $\frac{\partial T}{\partial t} = \frac{k}{\rho c} \frac{\partial^2 T}{\partial x^2}$ is $T = \exp(-x/\sqrt{at})$, which indicate that T is

nonzero at any long distance from the wall. Since the thermal wave propagates with a limited velocity (sound velocity), this is obviously impossible. To compensate for the error, the **Cattaneo equation** is introduced

$$\bar{\tau} \frac{\partial J_q}{\partial t} + J_q = -k \frac{\partial T}{\partial x}$$

where $\bar{\tau}$ is a weighed average of the relaxation time relative to the heat flux expression.

Combining this equation with the energy conservation equation (no heat generation considered),

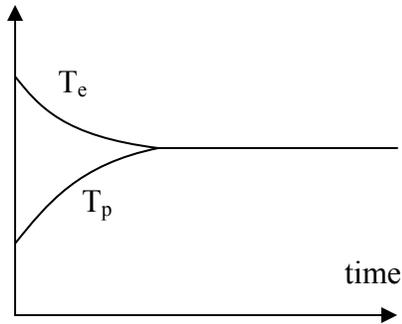
$$-\nabla \cdot \mathbf{J}_q = \rho c \frac{\partial T}{\partial t}$$

and eliminating J_q , we arrive at the following governing equation for the temperature distribution,

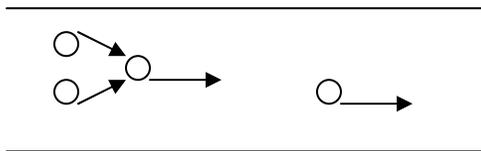
$$\bar{\tau} \frac{\partial^2 T}{\partial t^2} + \frac{\partial T}{\partial t} = \frac{k}{\rho c} \frac{\partial^2 T}{\partial x^2},$$

which is a **hyperbolic** type of equation.

However, the Cattaneo equation is not applicable in most experimental situations. Under fast heating, the temperature gradient is usually very large, and thus the condition that $(\Lambda/T \, dT/dx)$ cannot be satisfied. There is no convincing experimental data showing the validity of the hyperbolic equation. In femto-laser heating, the temperature of electrons is raised much higher than that of the phonons and after the relaxation time electrons exchange energy with phonons (the following figure).

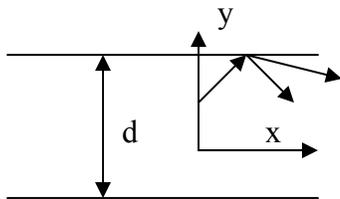


An experiment that is often cited as the proof for the validity of the hyperbolic equation is the second sound, which is encountered in low-temperature physics. However, the second sound is due to a different physical mechanism. At low temperatures, the umklap scattering is weak and normal scattering is strong, such that phonons have a nonzero average momentum (velocity). Equations for phonon hydrodynamics can be developed and the second sound can be described by these equations rather than the hyperbolic equation.



The derivations so far are for constitutive equations. We can also derive conservative equations from the Boltzmann equation. I will not go to details. Such conservative equations can be developed for gas molecules (Navier-Stokes equations), for electrons and phonons. Note: (1) In the 80s, people also conducted research based on Navier-Stokes type of equations for electron transport. However, it is not valid at nanoscale because just as Newton's shear stress law is not valid for rarefied gas flow, the drift-diffusion equation is not applicable to electrons.

Chapter 7 Classical Size Effects



When the electron and/or phonon mean free paths are comparable to or larger than the thin film thickness, they will collide more with the boundaries. The previous requirement $\Lambda/L \ll 1$ is not satisfied in this case. Most reflection on the wall is elastic. Two possibilities occur here: specular reflection, diffuse reflection.

First we have the Boltzmann equation

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f = -\frac{f - f_0}{\tau}$$

or

$$\mathbf{v} \cdot \nabla_{\mathbf{r}} f_0 + \mathbf{v} \cdot \nabla_{\mathbf{r}} g + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_0 + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} g = -\frac{g}{\tau}.$$

Ignore $\frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} g$. Magnitude comparison suggests that $v_y \frac{\partial g}{\partial y} \gg v_x \frac{\partial g}{\partial x}, v_z \frac{\partial g}{\partial z}$. Therefore,

we obtain

$$v_x \frac{df_0}{dx} + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_0 = -\frac{g}{\tau} - v_y \frac{\partial g}{\partial y},$$

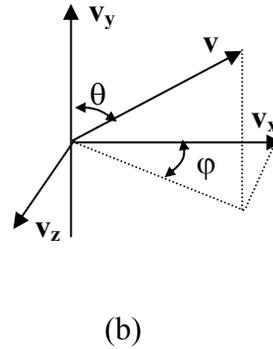
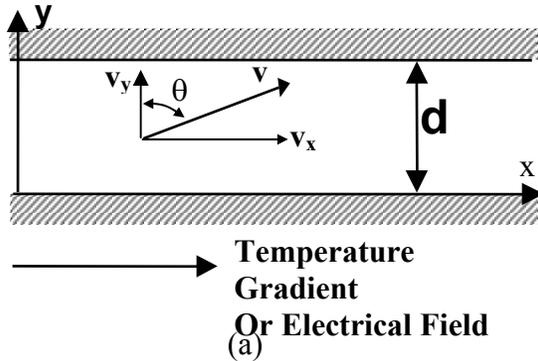
$$\tau v_x \frac{df_0}{dx} + \tau \frac{F_x}{m} \frac{df_0}{dv_x} = -g - \tau v_y \frac{\partial g}{\partial y}.$$

Note: For boundary scattering, the function $g(\mathbf{v})$ is unchanged but $g(\mathbf{r})$ is changed by scattering. For statistical calculation, caution needs to be taken when summation is conducted.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 22

We have talked about the heat flux as

$$q_x = \frac{1}{V} \sum_{k_x} \sum_{k_y} \sum_{k_z} f v_x \hbar \omega.$$



The Boltzmann equation is

$$\tau \vec{v} \cdot \nabla_{\vec{r}} g + g = -\tau \left(\vec{v} \cdot \nabla_{\vec{r}} f_0 + \frac{\vec{F}}{m} \cdot \nabla_{\vec{v}} f_0 \right) = S_0,$$

where $\frac{\vec{F}}{m} \cdot \nabla_{\vec{v}} f_0 = 0$ for phonons, $g = f - f_0$. Noticing $\vec{v} \cdot \nabla_{\vec{r}} g \approx v_y \frac{\partial g}{\partial y}$ ($d \ll x$) and

$\nabla_{\vec{r}} f_0 = \frac{df_0}{dT} \frac{dT}{dx}$, the x direction component gives

$$g + \tau v_y \frac{\partial g}{\partial y} = -\tau v_x \frac{df_0}{dT} \frac{dT}{dx} = S_0(x),$$

the solution of which is

$$g - S_0 = C \exp\left(-\frac{y}{v_y \tau}\right).$$

One boundary condition is required to determine C.

Assuming both top and bottom of the film diffusely scatter phonons, we have

$$\begin{cases} y = 0, f = f_0, g = 0 \text{ for } \theta \in \left(0, \frac{\pi}{2}\right) \\ y = d, f = f_0, g = 0 \text{ for } \theta \in \left(\frac{\pi}{2}, \pi\right) \end{cases}.$$

Finally we get

$$\text{At } y = 0, \theta \in \left(0, \frac{\pi}{2}\right), C = -S_0, g(y, \theta) = S_0 \left(1 - \exp\left(-\frac{y}{v\tau \cos \theta}\right)\right),$$

$$\text{At } y = d, \theta \in \left(\frac{\pi}{2}, \pi \right), C = -S_0 \exp(d / v\tau \cos \theta), g(y, \theta) = S_0 \left(1 - \exp\left(\frac{d-y}{v\tau \cos \theta} \right) \right).$$

Note: $\frac{y}{v\tau \cos \theta}$ is the ratio between the traveled distance and the mean free path of phonons.

$$q_x(y) = \frac{1}{V} \sum_{v_x} \sum_{v_y} \sum_{v_z} \hbar \omega v_x f = \int_0^{\omega_{\max}} d\omega \int_0^{2\pi} d\varphi \int_0^{\pi} \hbar \omega v_x f \frac{D(\omega)}{4\pi} \sin \theta d\theta,$$

where $v_x = v \sin \theta \cos \varphi$ according to our spherical coordinate system, $f = g + f_0$. Thus

$$q_x(y) = \int_0^{\omega_{\max}} d\omega \int_0^{2\pi} d\varphi \left[\int_0^{\frac{\pi}{2}} \hbar \omega (v \cos \varphi \sin \theta) \left(-\tau v \cos \varphi \sin \theta \frac{df_0}{dT} \frac{dT}{dx} \left(\exp\left(-\frac{y}{v\tau \cos \theta} \right) - 1 \right) \right) \frac{D(\omega)}{4\pi} \sin \theta d\theta + \int_{\frac{\pi}{2}}^{\pi} \hbar \omega (v \cos \varphi \sin \theta) \left(-\tau v \cos \varphi \sin \theta \frac{df_0}{dT} \frac{dT}{dx} \left(\exp\left(\frac{d-y}{v\tau \cos \theta} \right) - 1 \right) \right) \frac{D(\omega)}{4\pi} \sin \theta d\theta \right]$$

Note: At $y=0$, the x-direction heat flux is nonzero. In the above expression, the second integral is not zero though the first one is.

Now let $\cos \theta = \mu$ and thus $\sin \theta d\theta = -d\mu, \sin^3 \theta d\theta = -(1-\mu^2) d\mu$. Also define

$$E_n(\xi) = \int_0^1 \mu^{n-2} \exp\left(-\frac{\xi}{\mu}\right) d\mu, \quad \xi = \frac{d}{\Lambda}, \quad \Lambda = \tau v.$$

Noting $\int_0^{2\pi} \cos^2 \varphi d\varphi = \pi$, the total heat

flow per unit width is

$$Q = \int_0^d q(y) dy = -\frac{\pi}{4} \frac{dT}{dx} \int_0^{\omega_{\max}} \hbar \omega \frac{df_0}{dT} \tau v^2 D(\omega) d\omega \left[\int_0^1 (1-\mu^2) d\mu \left(\Lambda \mu \left(\exp\left(-\frac{\xi}{\mu}\right) - 1 \right) - d \right) + \right]$$

$$\int_0^1 (1-\mu^2) d\mu \left(-\Lambda \mu \left(1 - \exp\left(-\frac{\xi}{\mu}\right) \right) - d \right)]$$

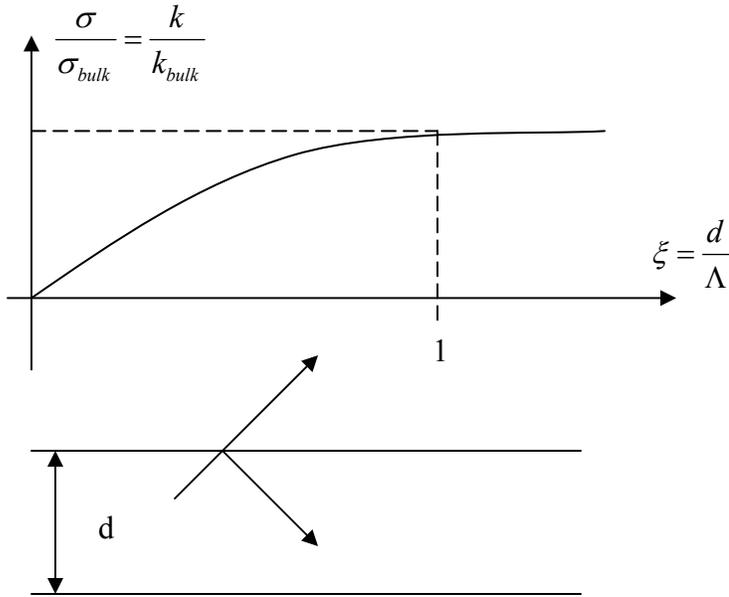
$$= -kd \frac{dT}{dx},$$

which yields (if Λ is independent on ω)

$$k / k_{\text{bulk}} = 1 - \frac{3}{8\xi} (1 - 4E_3(\xi) - E_5(\xi)).$$

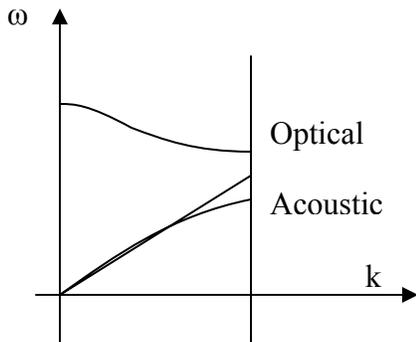
The tendency is drawn in the following figure. It is expected that k approaches the bulk value for d larger than the mean free path. Since in diffuse scattering part of the phonons

are scattered backward, the loss of x-direction momentum results in a lower thermal conductivity.



For specular case (above figure), the x-direction momentum is conserved. Following all these procedures, we can finally prove $k=k_{bulk}$.

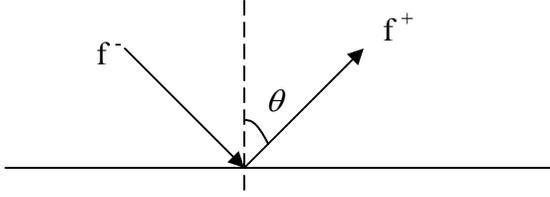
Note: To determine the mean free path, we should use $k = \frac{\bar{\Lambda}}{3} \int C_{\omega} v_{\omega} d\omega$ instead of the simplified $k = \frac{Cv\Lambda}{3}$, which gives an underestimation of Λ . This is because the Debye approximation overestimates the velocity approach the edge of the first Brillouin zone, where the group velocity should be zero.



For partial specular (momentum conserved) and partial diffuse (momentum not conserved), we have

$$f^+(\mu, 0) = pf^-(\mu, 0) + 2(1-p) \int_0^1 f^-(\mu, 0) \mu d\mu,$$

where p represents the ratio of specular scattering.



Now consider the y direction. Similarly, we have

$$\tau v_y \frac{\partial g}{\partial y} + g = -\tau v_y \frac{df_0}{dT} \frac{dT}{dy} = S_0(y).$$

To solve this, first let $S_0(y)=0$, then $g = C \exp\left(-\frac{y}{\tau v_y}\right)$. In general, use $C(y)$ to replace

C . Substitute into the governing equation, we get

$$\tau v_y \exp\left(-\frac{y}{\tau v_y}\right) \frac{dC}{dy} = S_0(y),$$

$$C = \int_y^{y_0} S_0(y') \frac{\exp\left(\frac{y'}{\tau v_y}\right)}{\tau v_y} dy' + C(y_0),$$

$$g(y) = C(y_0) \exp\left(-\frac{y}{\tau v \cos \theta}\right) + \int_y^{y_0} S_0(y') \frac{\exp\left(\frac{y'-y}{\tau v_y}\right)}{\tau v_y} dy'.$$

Now the boundary condition is (elastic scattering on boundaries)

$$\begin{cases} y=0, f=f_0, g=0, C(0)=0 \text{ for } \theta \in \left(0, \frac{\pi}{2}\right) \\ y=d, f=f_0, g=0, C(d)=0 \text{ for } \theta \in \left(\frac{\pi}{2}, \pi\right) \end{cases}$$

At steady state, we have

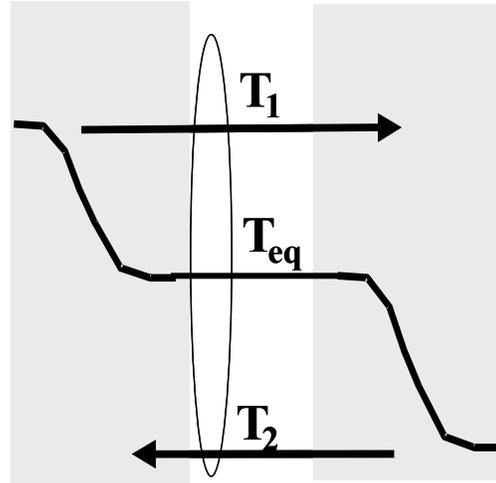
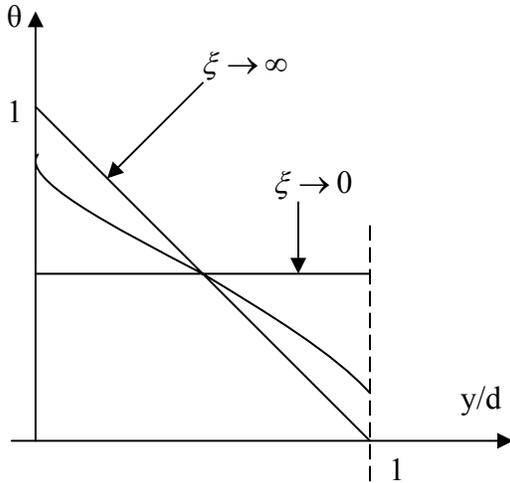
$$Const = q_y = \int_0^{\omega_{\max}} \hbar \omega d\omega \int_0^{2\pi} d\varphi \int_0^{\pi} v_y f \frac{D(\omega)}{4\pi} d\theta$$

And $\frac{dq_y}{dy} = 0$ yields

$$2\theta(\eta) = E_2(\eta) + \int_0^{\xi} \theta(\eta') E_1(|\eta - \eta'|) d\eta'$$

where we used dimensionless parameters $\theta(y) = \frac{T_1(y) - T_2}{T_1 - T_2}$, $\eta = \frac{y}{\Lambda}$, $\xi = \frac{d}{\Lambda}$.

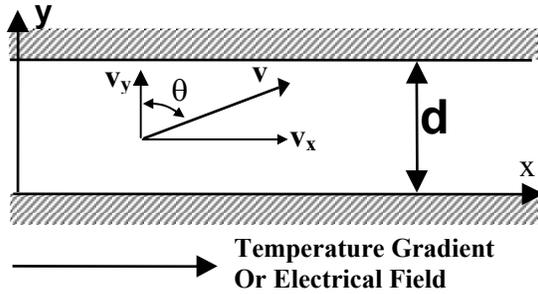
The temperature profiles for two extreme cases are drawn in the following figure. For $\xi \rightarrow 0$ (note $T_1 \neq T_2$), it is in nonequilibrium state but we define the equilibrium conception, temperature, based on the average value.



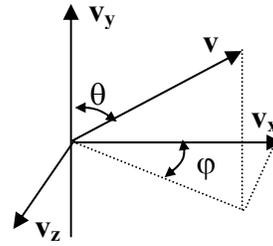
2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 23

In the last lecture, we talked about the energy flux along a thin film as

$$q_x = \frac{1}{V} \sum_{k_x} \sum_{k_y} \sum_{k_z} f v_x \hbar \omega.$$



(a)



(b)

Now let us consider the conduction of gas molecules between two plates. Similarly, the Boltzmann equation is

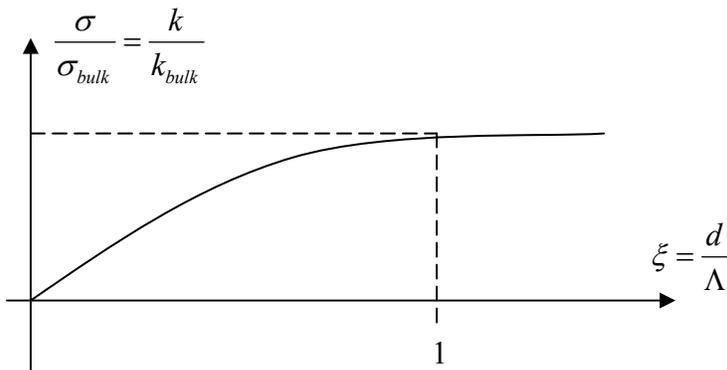
$$\tau \vec{v} \cdot \nabla_r \mathbf{g} + \mathbf{g} = -\tau \left(\vec{v} \cdot \nabla_r f_0 + \frac{\vec{F}}{m} \cdot \nabla_v f_0 \right),$$

where the bulk force $\frac{\vec{F}}{m} \cdot \nabla_v f_0 = 0$ for phonons, $\mathbf{g} = \mathbf{f} - f_0$. Noticing $\vec{v} \cdot \nabla_r \mathbf{g} \approx v_y \frac{\partial \mathbf{g}}{\partial y}$ ($d \ll x$)

and $\nabla_r f_0 = \frac{df_0}{dT} \frac{dT}{dx}$, the x direction component gives

$$g + \tau v_y \frac{\partial g}{\partial y} = -\tau v_x \frac{df_0}{dT} \frac{dT}{dx} = S_0(x).$$

We can solve g first and then substitute the expression $f = g + f_0$ into any flux equation. Under the diffuse assumption, we obtain the following figure for the conductivities of the material.



In the y direction, we have

$$\tau v_y \frac{\partial g}{\partial y} + g = -\tau v_y \frac{df_0}{dy}.$$

Since no heat is generated in the volume, we finally obtain $q(y) = \text{const}$. The

normalized temperature $\theta(\eta) = \frac{T(\eta) - T_2}{T_1 - T_2}$ obeys

$$2\theta(\eta) = E_2(\eta) + \int_0^\xi \theta(\eta') E_1(|\eta - \eta'|) d\eta',$$

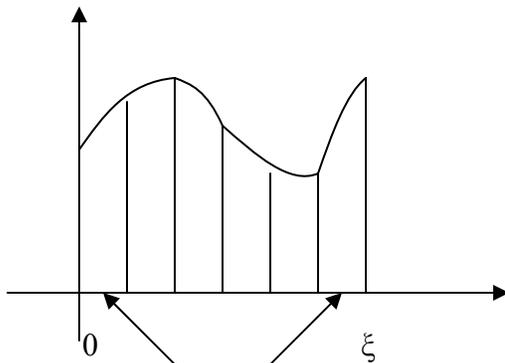
where $\eta = y/\lambda$, $\xi = d/\lambda$, $\lambda/d = \text{Knudsen number}$, $\theta(\eta)$ may also represent normalized

blackbody emissive power $\frac{u(y) - u_2}{u_1 - u_2} = \frac{T^4(y) - T_2^4}{T_1^4 - T_2^4}$. This is the linear, nonhomogeneous,

Fredholm integral equation of the second kind. The function $E_1(|\eta - \eta'|)$ is called the kernel. The Fredholm integral equation does not have an analytical solution, although approximate solution methods have been developed.

To solve the equation numerically, first we discretize the equation with

$$\frac{d^2 T}{dx^2} = \frac{T_{i+1} - 2T_i + T_{i-1}}{2\Delta x^2}.$$



Half trapezium at both ends

The integration is calculated by dividing the area into many trapezia. We have

$$\int_0^\xi \theta(\eta') E_1(|\eta - \eta'|) d\eta' = \theta(0) \frac{E_1(\eta)}{2} \Delta\eta + \theta(\xi) \frac{E_1(\xi - \eta)}{2} \Delta\eta + \Delta\eta \sum_1^{n-1} \theta(\eta_i) E_1(|\eta - \eta_i|),$$

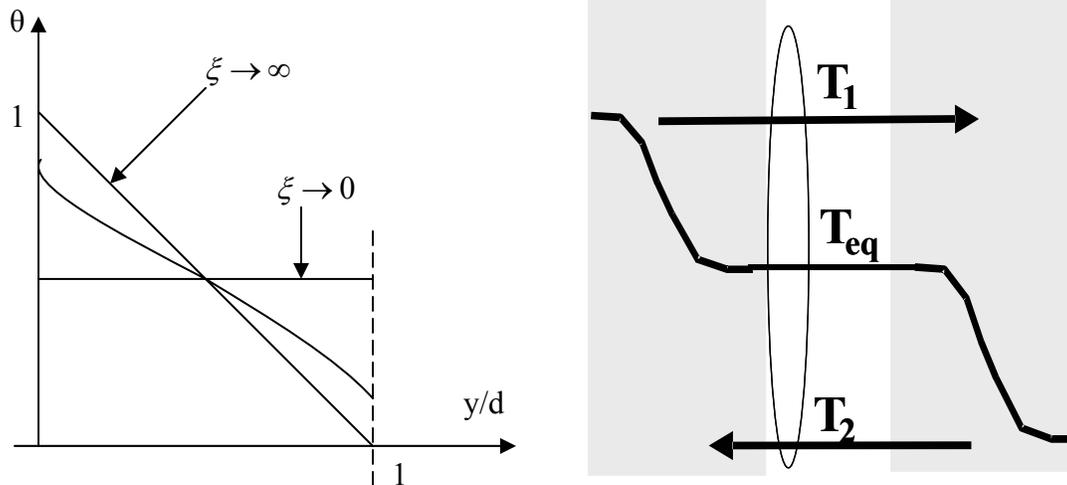
which gives totally $n+1$ equations (n is the section number). Therefore, we get a set of linear equations for $\theta(\eta_i)$ to solve.

Another way to conduct the integration is to use the Gauss- Legendre method, which

generates sections in $(0, \xi)$ with varying width. Then we can use $\int_0^1 f(x) dx = \sum_{i=1}^n f(x_i) W_i$

to find the approximating value, where W_i is the weight of different $f(x_i)$.

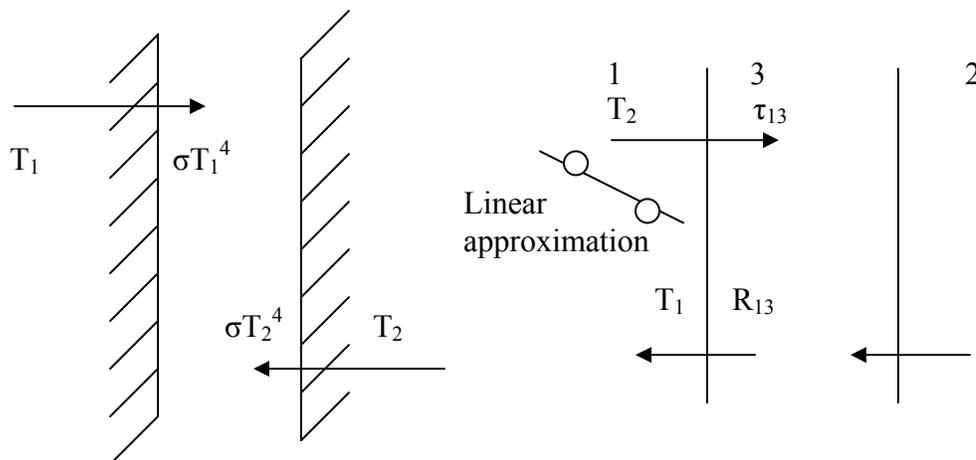
The temperature $\theta(\eta) = \frac{T(\eta) - T_2}{T_1 - T_2}$ within d is shown as following, which was discussed in the last lecture. When $\xi = d/\lambda \rightarrow 0$, $T(y) = \frac{T_1 + T_2}{2}$.



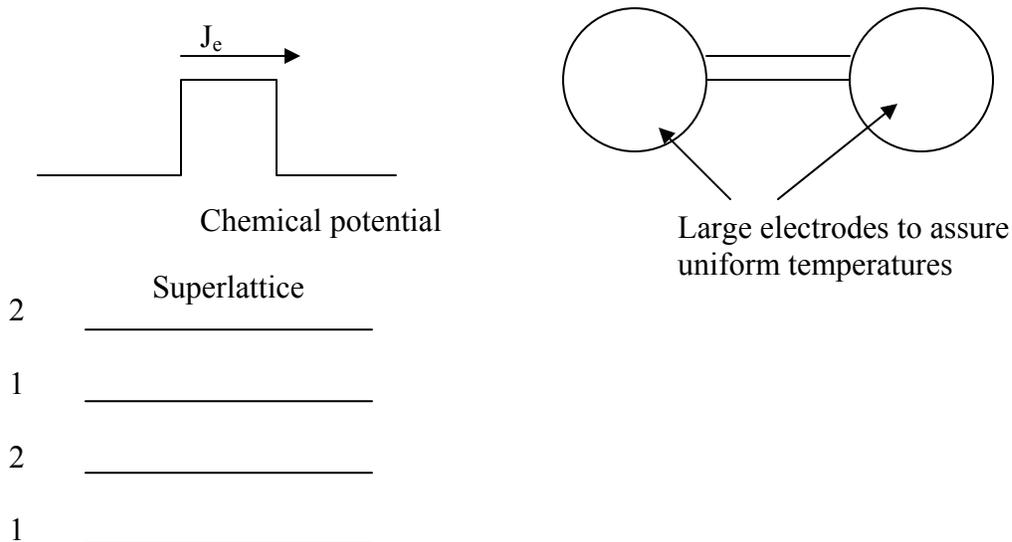
More generally, θ should be the internal energy of the carriers. For photons,

$$\theta(y) = \frac{u(y) - u_2}{u_1 - u_2} = \frac{T^4(y) - T_2^4}{T_1^4 - T_2^4}.$$

The interpretation of the temperature discontinuity is worthy of special attentions. Sometimes, the jump is physical, while in other cases the jump is artificial. The discontinuity arises from the boundary conditions. We assume the boundaries are diffuse and carriers coming into the region are at temperatures T_1 and T_2 (emitted temperatures). These temperatures represent only those carriers entering the thin film. For photon transport, these are good representations of the carriers coming into the region because the internal conduction in both walls is much stronger and determines the solid temperatures that emit the photons. In this case, the temperature jump represents the difference of the solid temperature and the photon temperature inside the film.

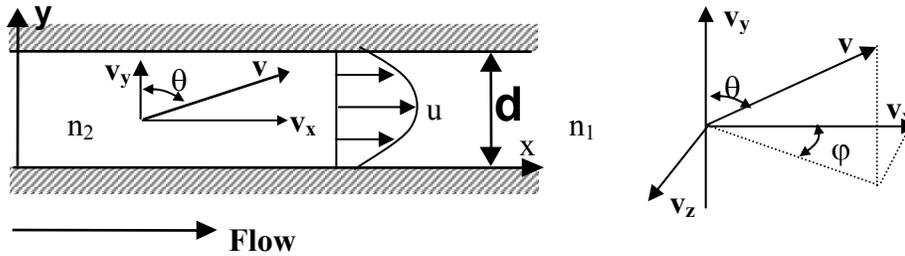


For phonons, if the walls are assumed to be black, it also means that transmissivity equals one and there exists no thermal boundary resistance. The phonon temperature inside region 1 should equal to that temperature in the film at the boundary. The temperature jump is artificial, and arises from the fact T_1 and T_2 are only emitted phonon temperature entering the thin film, not the local equilibrium temperature as we use in the Fourier law or solved directly from the Boltzmann equation. However, if the interface reflectivity is not zero, there exists a temperature jump just as in the case of thermal boundary resistance that we discussed in chapter 5. The value of this temperature jump is different from what is given from the boundary condition used. This issue related to consistent use of a definition for temperature was also discussed for the case of thermal boundary resistance. Many comparisons between experimental data and theory are based on inconsistent definition on temperature and are thus wrong (theoretically). In heat conduction experiments, one usually measure two points close to the interface and use linear approximation to get the temperatures on both sides of the interface. The temperature obtained is consistent with the Fourier law but not consistent with the boundary conditions we use here in the solution of the Boltzmann equation. Iterative procedures are needed if one want to modify the boundary condition to be consistent with the temperatures defined in the Boltzmann equation and the Fourier law. Similar phenomena also happen in when dealing with electron transport across an interface. Most people use drift-diffusion equations up to the interface and use Richardson formula (which is based on the emitted electron properties including their chemical potential and temperature). There are cases, however, one measure the emitted electron and phonon temperature or chemical potential, as in the case of photon radiation discussed above. One example is in the experiment determining the quantized conductance of a nanowire. The electrodes are large and the measured voltage drop should represent the differences of electrons entering the channel.

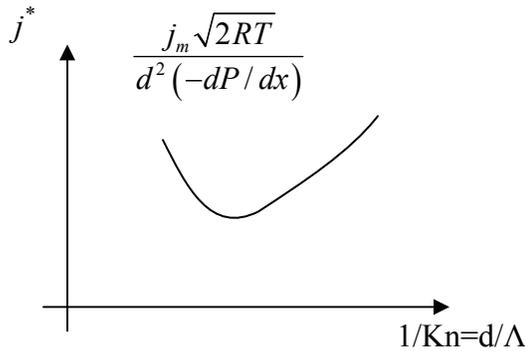


I found the importance of a consistent usage of the temperature definition when dealing with heat conduction in superlattices. For the above superlattice case, a consistent

temperature definition must be used such that results from Boltzmann equation and from simple ray tracing in the limit of no scattering agrees with each other.

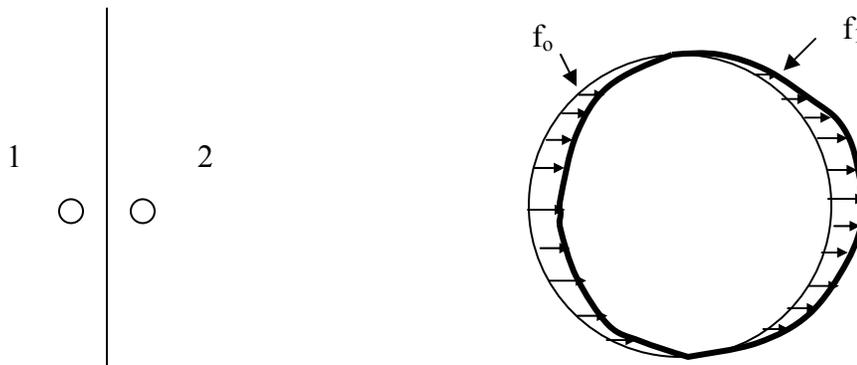


For gas flow through two plates (driven by pressure difference), the velocity has a drop on the surface, which is called slip boundary condition. The mass flow rate per unit depth is drawn in the following figure. The mass flow rate demonstrates the existence of the Knudsen minimum around $K_n=1$ and is in reasonably good agreement with experimental results of Dong (1956).



Now let us consider the interface temperature drop with learned equations. Here we use diffusion approximation inside ($\vec{v} \cdot \nabla_{\vec{r}} g$ is dropped in Boltzmann equation), and assume diffusion/transmission boundary conditions. The Boltzmann equation becomes

$$f - f_0 = -\tau \left(\vec{v} \cdot \nabla_{\vec{r}} f_0 + \frac{\vec{F}}{m} \cdot \nabla_{\vec{v}} f_0 \right)$$



In the above left figure, at the interface we have

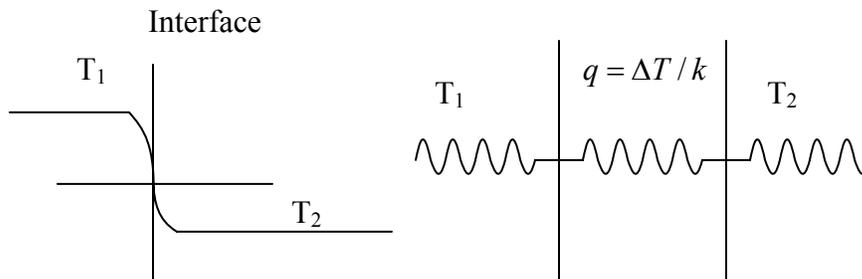
$$q = \sum_{v_{x1} > 0, v_{y1}, v_{z1}} \tau_{12} \hbar \omega v_{x1} f_1 + \sum_{v_{x2} > 0, v_{y2}, v_{z2}} \tau_{21} \hbar \omega v_{x2} f_2.$$

Note in the right figure, only rightward arrows indicate transport to the second region ($V_x > 0$). Define τ_{12}' , τ_{21}' as the average transmissivity in each region. We have

$$q = \sum_{v_{x1} > 0, v_{y1}, v_{z1}} \tau_{12} \hbar \omega v_{x1} f_{01} + \sum_{v_{x2} > 0, v_{y2}, v_{z2}} \tau_{21} \hbar \omega v_{x2} f_{02} + \tau_{21}' \frac{q}{2} + \tau_{12}' \frac{q}{2},$$

$$\left[1 - \frac{1}{2}(\tau_{21}' + \tau_{12}')\right]q = \sum_{v_{x1} > 0, v_{y1}, v_{z1}} \tau_{12} \hbar \omega v_{x1} [f_{01}(T_1) - f_{02}(T_2)].$$

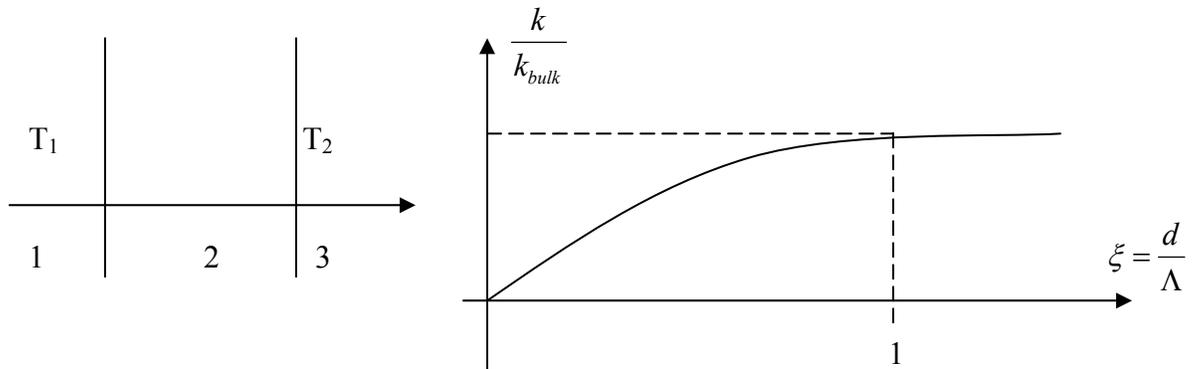
If $\tau_{21}' = \tau_{12}' = 1$, no temperature drop happens in the above relationship. In practice, τ_{21}' , τ_{12}' are difficult to determine.



For the case drawing as below, we have

$$\frac{k_{eff}}{k_{bulk}} = \frac{3d/4\Lambda}{\frac{1 - \frac{1}{2}(\tau_{12}' + \tau_{21}')}{\tau_{21}''} + \frac{1 - \frac{1}{2}(\tau_{23}' + \tau_{22}')}{\tau_{23}''} + 3d/4\Lambda},$$

where $3d/4\Lambda \ll \frac{1 - \frac{1}{2}(\tau_{12}' + \tau_{21}')}{\tau_{21}'} + \frac{1 - \frac{1}{2}(\tau_{23}' + \tau_{22}')}{\tau_{23}'}$ in thin films, while in bulk materials $3d/4\Lambda$ is dominant. We have similar relationship as previous cases.



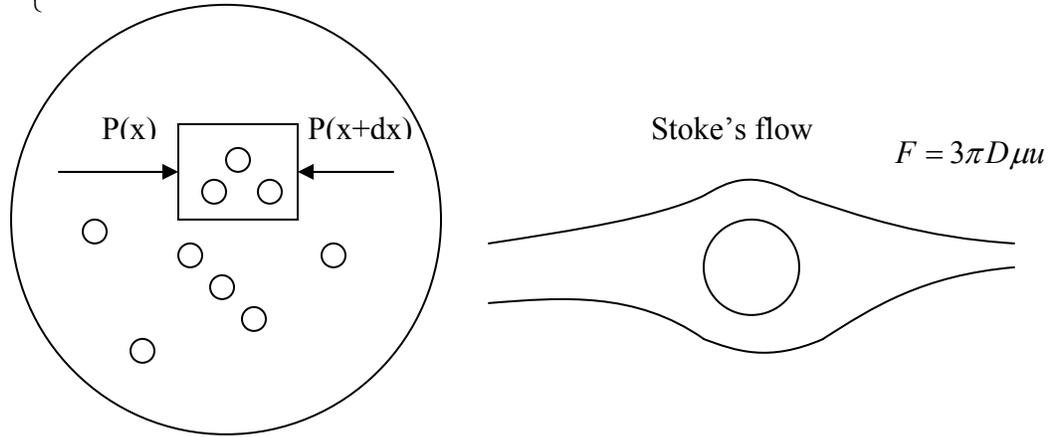
Chapter 9 Liquids

For gases, after two particles collide, they do not have any memory of previous history. However, this is not true for solids and liquids, in which their previous locations affect

the shape change. The Boltzmann equation is only applicable to diluted gases. To deal with liquids, other approaches must be used. We list some of the efforts in this aspect.

Approaches {

- Modified Boltzmann equation {
 - Two-particle distribution function
 - Enskog equation (denser gases)
- Einstein (Brownian motion) \Rightarrow Langerin equation \Rightarrow
- Linear response theory \Rightarrow Molecular Dynamics



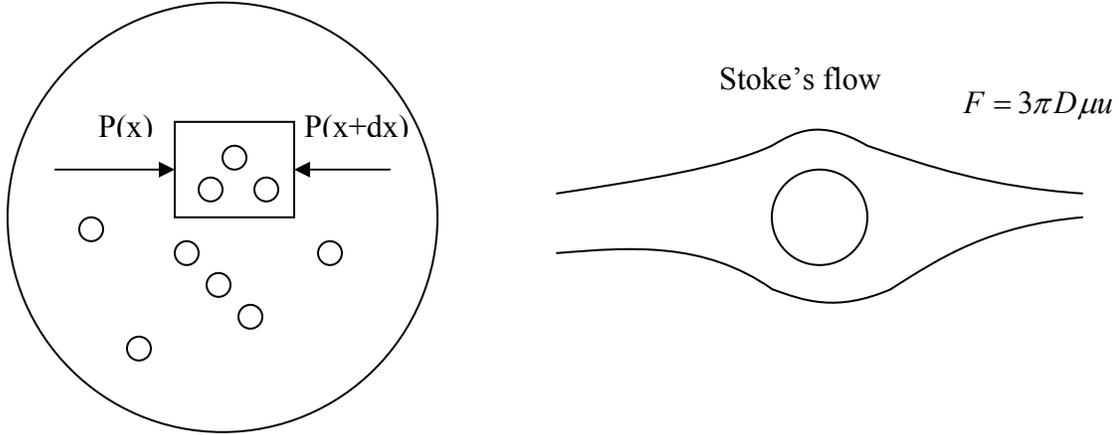
In the left figure above, pressure difference exists in the fluid. For one particle, the osmotic pressure is determined by

$$P = \frac{1}{V} Nk_B T = nk_B T .$$

Also, we can use the solution of Stoke's flow around a sphere. The force is $F = 3\pi D\mu u$.

2.57 Nano-to-Macro Transport Processes
Fall 2004
Lecture 24

In the last lecture, we have talked about Einstein's work on the Brownian motion.



In the left figure above, pressure difference exists in the fluid. For one particle, the osmotic pressure is determined by

$$P = \frac{1}{V} Nk_B T = nk_B T .$$

Also, we can use the solution of Stoke's flow around a sphere to estimate the drag force. The value is $F = 3\pi D\mu u$. For area A_c , the forced balance over the control volume gives

$$A_c (P(x) - P(x + dx)) = FdN ,$$

$$-A_c dx \cdot \frac{dP}{dx} - 3\pi D\mu u dN = 0 ,$$

$$-\frac{dP}{dx} - 3\pi D\mu un = 0 ,$$

$$-k_B T \frac{dn}{dx} - 3\pi D\mu un = 0 ,$$

in which the product un indicates flux. We have

$$J_p = -\frac{k_B T}{3\pi D\mu} \frac{dn}{dx} = -a \frac{dn}{dx} ,$$

where $a = \frac{k_B T}{3\pi D\mu}$ is the diffusivity and can be obtained by the diffusion experiment of

some materials. In a time t , the diffusing radius is $\Delta r = \sqrt{6at}$. We can calculate the constant a and substitute back into $a = \frac{k_B T}{3\pi D\mu}$, thus D can be obtained. The relationship

between thermal diffusivity and viscosity is an example of the more general fluctuation-dissipation theory as viscosity is a measure of dissipative process and diffusivity is a measure of random walk (fluctuation) process. The relationship between thermal

diffusivity and viscosity is also called Einstein relation. In chapter 6, there is a similar Einstein relation between electron diffusivity and mobility. But Einstein really worked on the Brownian motion.

Note: The osmotic pressure can be observed by putting a semi-permeable membrane such that only the base molecules in the solution penetrate the membrane, building up a concentration gradient on the two sides of the membranes.

In addition to the above approach, Einstein also established another method to determine the diameter of the solute particles. He proposed to measure the viscosity of the solvent and of the solution, μ , and μ_0 , respectively, and derived, again assuming dilute solute particles, the following relationship between the viscosities

$$\frac{\mu}{\mu_0} = 1 + 2.5\varphi,$$

where φ is the volumetric concentration of the solute molecules.

The Einstein relation can also be derived from the stochastic approach developed by Langevin to treat Brownian motion of particles much larger than that of the surrounding medium. The key idea of the Langevin equation is to consider that the motion of a Brownian particle is subject to a friction force that is linearly proportional to its velocity, as in Stokes law, and a random driving force (or “noise”), $\mathbf{R}(t)$, imparted by the random motion of the molecules in the bath. In the absence of an external force, the Langevin equation that governs the instantaneous velocity of the Brownian particle can be written as,

$$m \frac{d\mathbf{u}}{dt} = -m\eta\mathbf{u} + \mathbf{R}(t),$$

where η is the friction coefficient, and for Brownian particles in a fluid the Stokes law gives $F = 3\pi D\mu u$, so that the random driving force $\mathbf{R}(t)$ has the following characteristics:

$$\langle \mathbf{R}(t) \rangle = 0 \quad (\text{average of random driving force is zero})$$

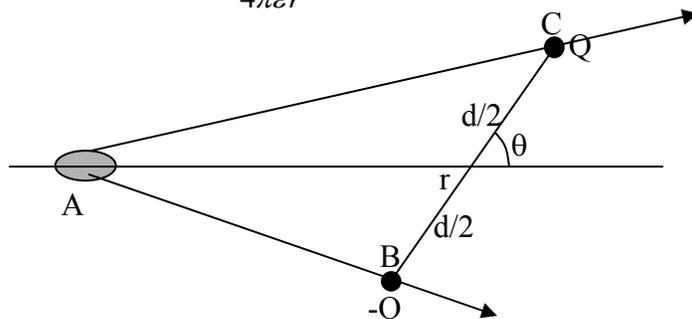
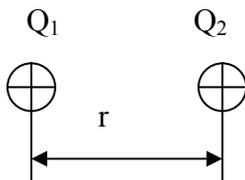
$$\langle \mathbf{R}(t) \cdot \mathbf{u}(t) \rangle = 0 \quad (\text{random driving force is not correlated to the velocity})$$

$$\langle \mathbf{R}(t+s) \cdot \mathbf{R}(s) \rangle = 2\pi R_0 \delta(t).$$

9.2 Force and potentials

For liquids, potential interaction contributes significant to transport. In previous lectures, we have discussed force $\vec{F} = -\nabla\Phi$. For the interaction of two charged particles as the

following left figure, the Coulomb potential is $\Phi = \frac{Q_1 Q_2}{4\pi\epsilon r}$, in which $\epsilon = \epsilon_{\text{vacuum}} \epsilon_r$.



You can build up all potentials from the Coulomb potential. Now let us consider the interaction of a charge and a dipole as shown in the right figure above case. We have

$$\phi = -\frac{Qq}{4\pi\epsilon_0} \left[\frac{1}{AB} - \frac{1}{AC} \right],$$

where

$$AB = \left[\left(r - \frac{d}{2} \cos \theta \right)^2 + \left(\frac{d}{2} \sin \theta \right)^2 \right]^{1/2},$$

$$AC = \left[\left(r + \frac{d}{2} \cos \theta \right)^2 + \left(\frac{d}{2} \sin \theta \right)^2 \right]^{1/2}.$$

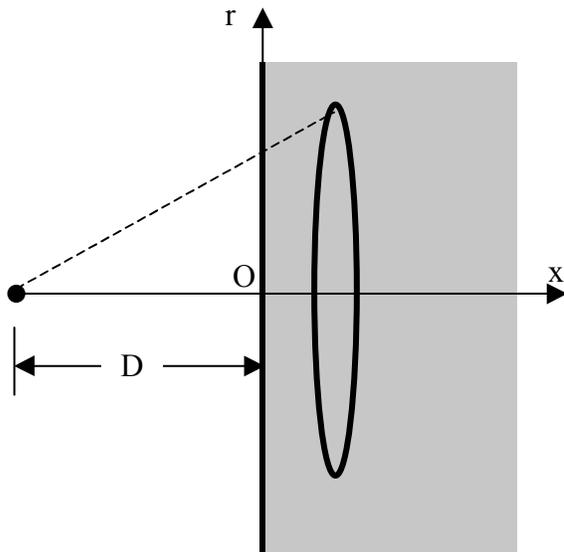
Under the approximation $r \gg d$, we obtain

$$\phi(r, \theta) = -\frac{q\beta \cos \theta}{4\pi\epsilon_0 r^2},$$

where $\beta = Qd$ is the dipole moment.

Note: (1) The superimposed two fields yield $\Phi \sim r^{-2}$ instead of $\Phi \sim r^{-1}$. (2) If we have two dipoles, the potential becomes $\Phi = -Cr^{-6}$.

The van der Waals potential between one atom and a surface

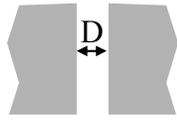


Based on the elementary potential interactions discussed in the previous section, the force interaction between particles and surfaces can be obtained by summing up the interactions between the atoms or molecules involved, as demonstrated in the above

figure. For two opposite surfaces, the Van der Waals force is $-\frac{A}{12\pi D^2}$. The **Hamaker constant**, after Hamaker (1937)

$$A = \pi^2 C n_1 n_2,$$

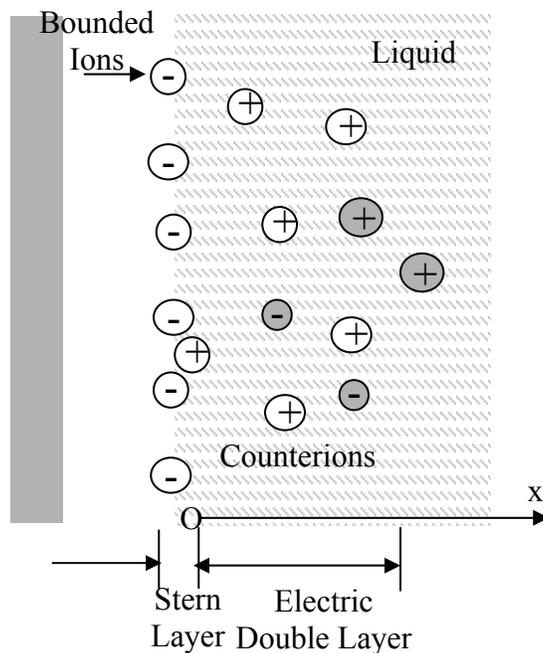
where n_1 and n_2 are the number density of molecules of the two interacting media.



$$\Phi = -\frac{A}{12\pi D^2}$$

Note: In $\vec{F} = -\nabla\Phi$, a positive A leads to positive \vec{F} (attractive force); otherwise it is repulsive force.

9.2.3 Electric double layer potential



(Solid surface should not be separated from the negative ions.)

Surfaces immersed in liquids are usually charged due to the ionization or dissociation of surface groups or due to the adsorption of ions from the solution onto a previously uncharged surface. The charges accumulated at the surface are balanced by an equal but oppositely charged region of counterions. Some of these counterions are also bounded to the surface, forming a so-called **Stern or Helmholtz layer**, which is usually very thin (a few Angstroms). The remaining counterions distribute near the surfaces but are free to move, forming a diffuse **electric double layer**. This electric double layer is of fundamental importance for a wide range of technologies such as batteries, fuel cells, colloids, and in biochemistry and biotechnology.

Note: You may compare this case to a p-n junction with the built-in potential. We first determine the magnitude of the potential developed on the solid-liquid interface. This potential can be easily measured using the solid as an electrode. The ion density on the solid-surface obeys the Boltzmann distribution,

$$c_s = c_{zp} e^{-Ze\psi_s / \kappa_B T},$$

where Z is the number of charges per ions and e is the unit charge ($-e$ for an electron), ψ_s is the electrostatic potential of the solid surface, and c_{zp} is the ion density at zero surface electrostatic potential. The above equation is also called the Nernst equation.

The Maxwell equation determines the displacement \bar{D} as $\nabla \cdot \bar{D} = \rho_{net}$, $\bar{D} = \epsilon \bar{E} = \epsilon (-\nabla \Psi)$.

Thus we obtain the **Poisson-Boltzmann equation**

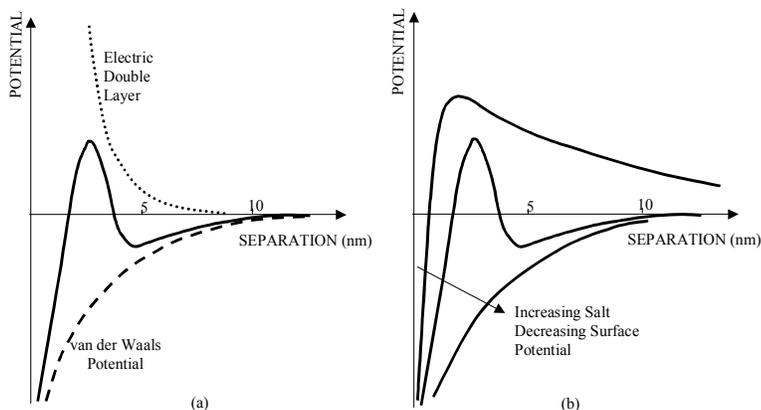
$$-\epsilon_0 \epsilon_r \frac{d^2 \Psi}{dx^2} = \rho_{net} = \sum_i Z_i e n_{oi} \exp\left(-\frac{Z_i e \psi}{\kappa_B T}\right).$$

Finally we obtain the **Debye length**,

$$\frac{1}{\delta} = \sqrt{\sum_i \frac{Z_i^2 e^2 n_{oi}}{\epsilon_0 \epsilon_r \kappa_B T}}.$$

Note: $\epsilon_r \uparrow, \delta \uparrow$.

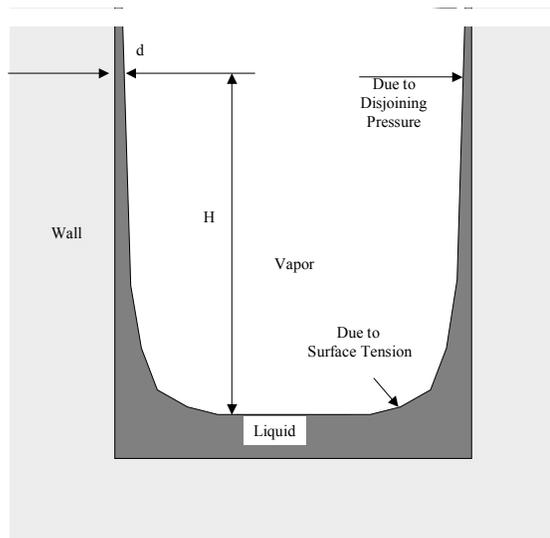
Now we consider the force balance inside the liquid. Because the liquid is stationary, the electrostatic force on liquid must balance the pressure force. The pressure inside the electric double layer is higher than that inside the bulk liquid at the equilibrium state. Furthermore, the electric double layer creates an attraction force between the ions on the solid surface and the counterions in the liquid. This attraction electrostatic force is balanced by the positive pressure in the liquid. When two solid surfaces are brought close to each other, a repulsive force develops between the two surfaces because the electrostatic force no longer balances the positive pressure inside the liquid.



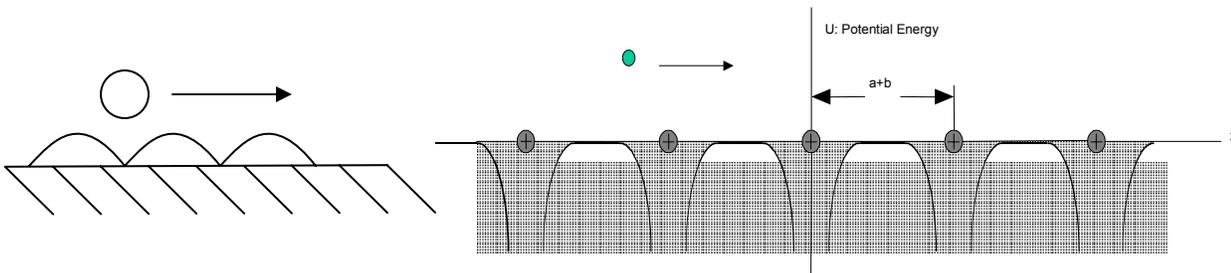
Now consider particles in a solution. Two main potential interactions exist. One is the van der Waals (usually attractive) and the other is double layer interaction (repulsive). In

the above figure (a), a superposition of the double layer potential (repulsive electro potential) and van der Waals potential (attractive) is shown. Under varying salt concentration, different combinations can be obtained (figure (b)). A colloid forms when there exists a minimum in potential so that particles are separated at a distance.

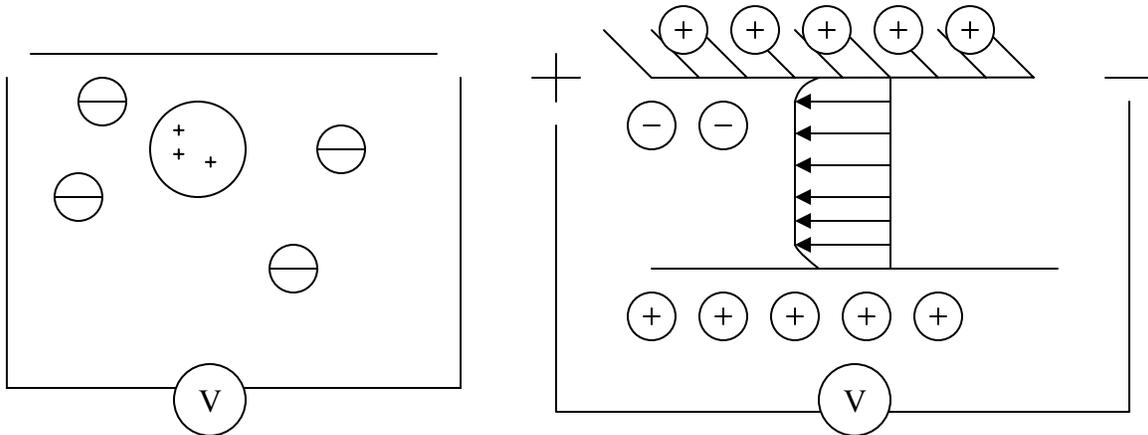
The repulse forces between surfaces also created another interesting phenomenon, the disjoining pressure. The following figure shows an example where the disjoining pressure plays an important role. At the base of liquid surface intersection the wall, the liquid layer climb up the wall due to surface tension. The vapor and the solid wall are two parallel surfaces. The medium in between is liquid. When the interaction force between wall and vapor is repulsive (due to a negative Hamaker constant or electric double layer), a positive pressure in the liquid layer develops, superimposed on normal compressive pressure. The thinner is the film, the more positive is the pressure. Which means that a lifting force (due to larger compression pressure) at the base will push liquid up the wall, till it is balanced by the gravitational force.



The appreciation of importance of potential and force interactions also allows one to do some quick orders of magnitude analysis. Can liquid slip on a wall. My conclusion, based on a simple orders of magnitude analysis, as given in an example in the book, says not in practical cases. You can analyze this problem by considering how much bonding force an atom experiences from the wall, and compare that to the shear stress generated in practical situations.



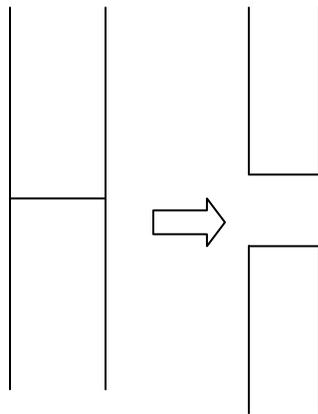
Electric double layer is the basis of many microfluidic technologies. Because the solid and the liquids are charged, an external field can create their motion. There are two cases, one is the particle in a liquid moves (called electrophoretic motion). This is the basis in DNA separation. In the following left figure, charged DNA molecules can be separated by the electric field. In the second case, the solid walls are stationary. The ions will move under an external electric field (called electrokinetic flow). The moving ions will also cause the fluid to move, and form a plug flow (flat profile). This is used for pumping purpose.



Surface tension

The idea of surface tension can be understood in terms of the energy stored in the surface. When we stretch the rod till it breaks, the work input is stored in the two new surfaces,

$$\text{i.e., } \gamma_1 = \frac{1}{2} W_{11}.$$



Now consider the separation of two immiscible liquids in contact into two stand-alone parts at the interface. After separation, the interfacial energy on each surface is γ_1 and γ_2 . The energy difference between the surface energy after the separation and the interfacial tension γ_{12} before the separation is called the work of adhesion,

$$W_{12} = \gamma_1 + \gamma_2 - \gamma_{12}.$$

The work of adhesion can be approximately estimated from the work of cohesion,

$$W_{12} = \sqrt{W_{11d}W_{22d}} = 2\sqrt{\gamma_{1d}\gamma_{2d}} .$$

Thus

$$\gamma_{12} = \gamma_1 + \gamma_2 - 2\sqrt{\gamma_{1d}\gamma_{2d}} .$$

Surface tension is very important for microsystems. There are two basic equations for analyzing surface tension. One is the **Laplace equation**

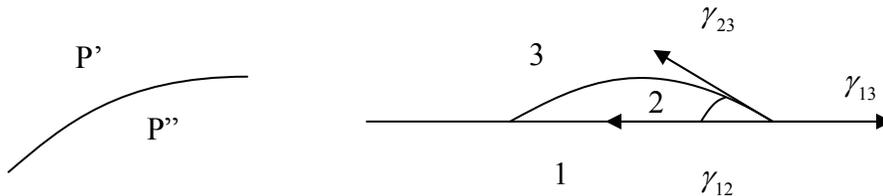
$$p'' - p' = \frac{2\gamma}{r} ,$$

if the two radii of curvature in two orthogonal directions are both r . The equation says there is a pressure difference between inside and outside a particle due to surface tension.

The other is for the interaction of three surfaces. For the following right case, the **Young equation** gives

$$\gamma_{13} = \gamma_{23} \cos \theta + \gamma_{12} .$$

The forces are shown in the figure.



Size can affect the phase change processes. One can easily appreciate this from the Laplace equation. Since pressures inside and outside a particle (droplet, bubble, solid particle) are different, and they are further related to other thermodynamic properties, it is reasonable to anticipate that certain thermodynamic properties will be influenced by the size (saturation temperature, pressure, enthalpy, surface tension). Due to the surface tension, an isolated nanoparticle usually has lower melting points compared with bulk materials. The equilibrium vapor pressure increases as the liquid droplet radius decreases. For a given vapor pressure, smaller droplets tend to evaporate.

MIT OpenCourseWare
<http://ocw.mit.edu>

2.57 / 2.570 Nano-to-Macro Transport Processes
Spring 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.