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 $\varphi = 2\pi n d \cos\theta / \lambda_0$. 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

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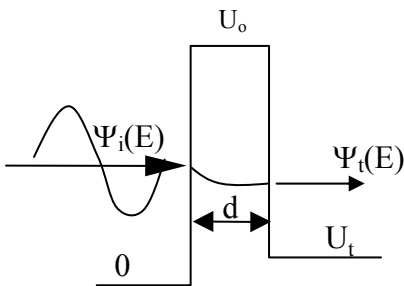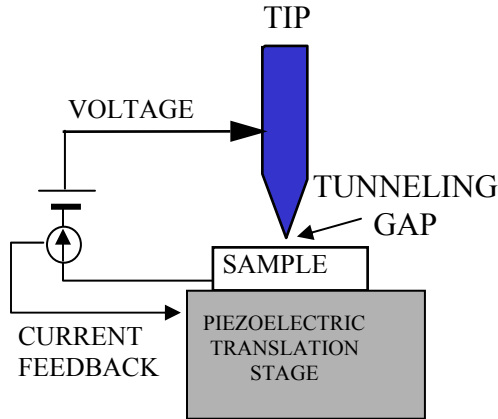

Now let us estimate the d value for a 1eV energy barrier. The k_2 value is

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

and the characteristic length for d is 1/k~1 Å.

Applications of tunneling
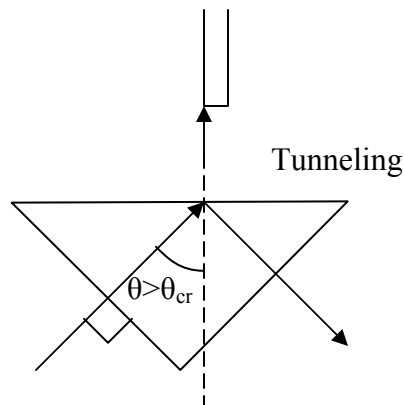1) Scanning tunneling electron microscope (STM)

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 Rhorer, 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 ~1 Å^{-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.
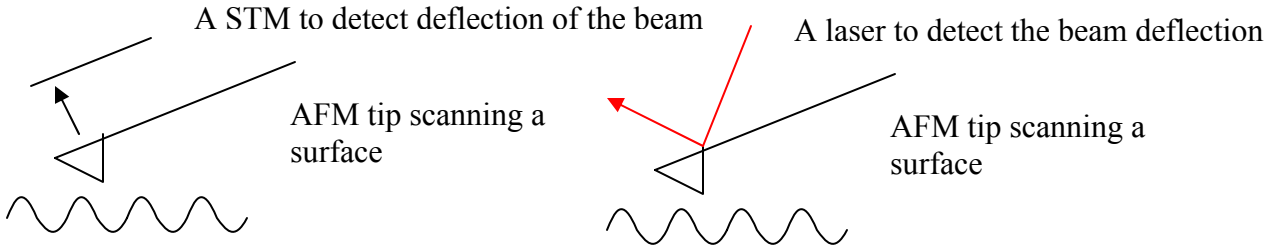
2) Other microscopes
A) Photon scanning tunneling microscope
In such a microscope, the incident angle on the scanned surface is larger than the critical angle and thus it will be totally reflected in normal cases. However, when a scanning probe approaches the other side of the surface, there will be 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.

A STM to detect deflection of the beam

A laser to detect the beam deflection

AFM tip scanning a surface

AFM tip scanning a surface

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.

Large effective contact area
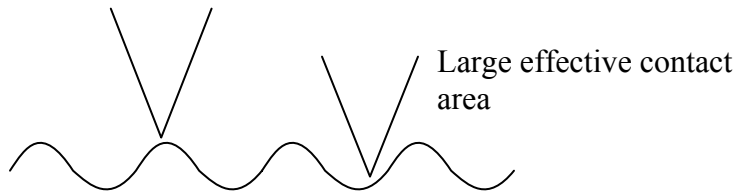
5.4 Bragg reflector

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 $2\ell\pi$.
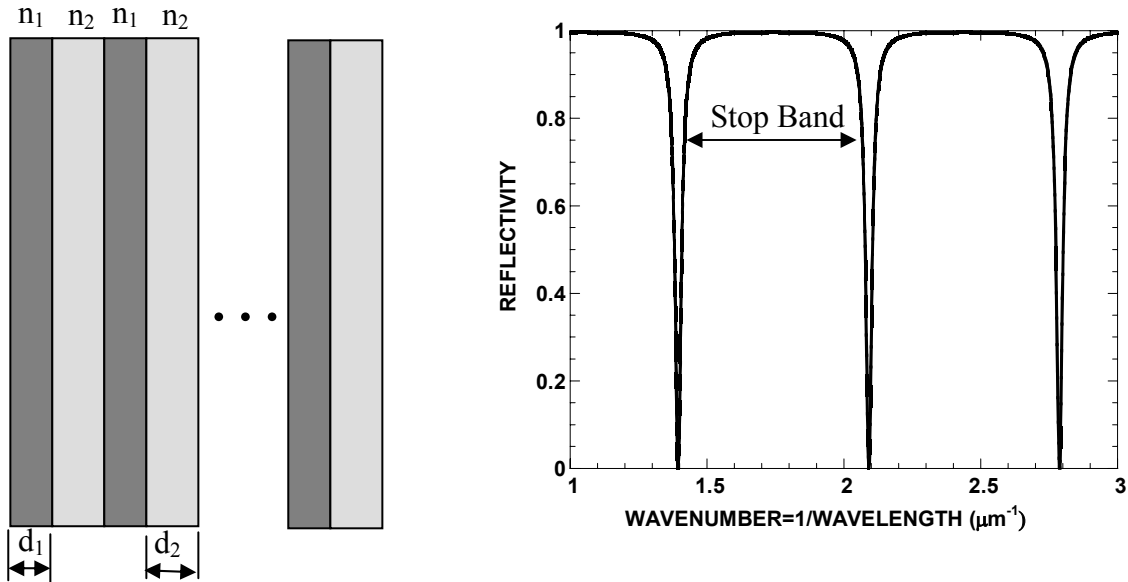
It is interesting to compare the stop band with electron band gap in semiconductors. In the latter case, no electrons exist in the forbidden energy levels, while in the reflector no photons exist in the stop band. The idea of forbidden energy in semiconductors is also utilized to filter electrons in superlattice, whose inventor, Esaki won a Nobel Prize for his work on double-barrier tunneling effect.

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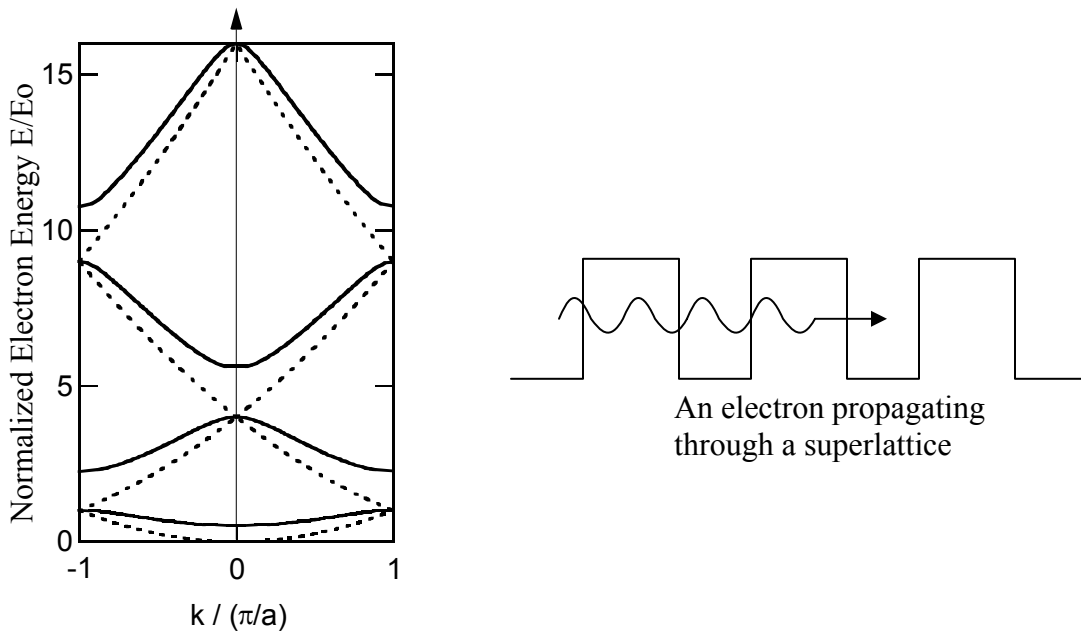
Note: The research in 3D semiconductors triggered the work on 1D superlattice, and now the research go back to 3D cases in 3D Bragg reflector, also called photonic crystals. Photonic band is formed within such a crystal and filter out radiation with unwanted wavelengths.

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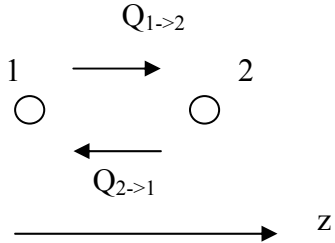
Now let us go back to the energy exchange between two points, which is expressed as

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

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.

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

in which

$$\frac{dk_{x1} dk_{y1} dk_{z1}}{(2\pi/L)^3} \frac{1}{V_1} = \frac{4\pi k^2 dk}{8\pi^3} = \frac{4\pi k^2}{8\pi^3} \frac{dk}{dE} dE = D(E_1) dE.$$
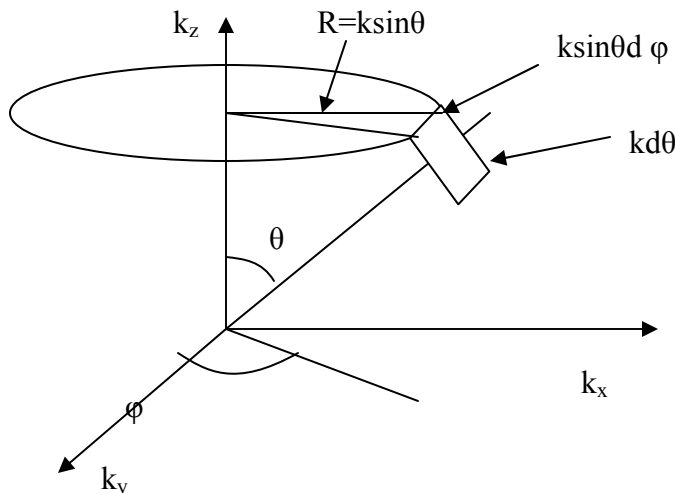
However, we also notice that velocity v_{z1} depends on k_{x1} and transmissivity τ_{12} depend on $\left(k_{x1}, k_{y1}, k_{z1}\right)$. To deal with this, first we note

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

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

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

For isotropic materials, $\dfrac{k^2}{8\pi^3}\dfrac{dk}{dE} = \dfrac{D(E)}{4\pi} = D'(E)$ because the solid angle of whole

sphere is 4π (or $\dfrac{D(E)}{4\pi} d\Omega dE = D(E)dE$).

The total energy flux is defined as
$$q_{12} = q_{1\to 2} - q_{2\to 1} = q_{1\to 2}\left(T_{e1}\right) - q_{2\to 1}\left(T_{e2}\right) = q_{1\to 2}\left(T_{e1}\right) - q_{1\to 2}\left(T_{e2}\right),$$
where we use $q_{1\to 2}\left(T_{e2}\right) = q_{2\to 1}\left(T_{e2}\right)$ 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 $q_{2\to 1}\left(T_{e2}\right)$. This idea is comparable to the Kirchoff's Law in radiation, in which we consider the energy exchange with a blackbody surface to derive $\varepsilon(\lambda,\theta) = \alpha(\lambda,\theta)$.
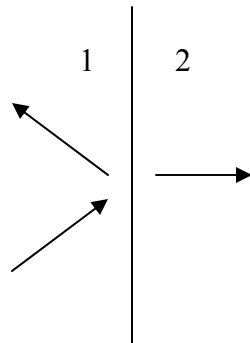
From above equations, we finally obtain the Landauer formalism
$$q_{1\to 2} = \sum_{p=1}^{3m}\left[\int_S d\Omega \int v_{z1}\, E_1 \tau_{12}\left[f\left(E_1, T_{e1}\right) - f\left(E_1, T_{e2}\right)\right]\frac{D(E_1)}{4\pi} dE\right],$$
where the solid angle is integrated over the hemisphere.

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 (W/m^2 K), R_{12} is the contact thermal resistance (m^2 K/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~v_{z1}C~1e3 $\times$ 1e6=1e9 W/m^2 K. (2) Based on one-dimensional heat conduction, $R = \frac{L}{kA}$ leads to the equivalent thickness of the interface is 0.1 μm 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 if 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
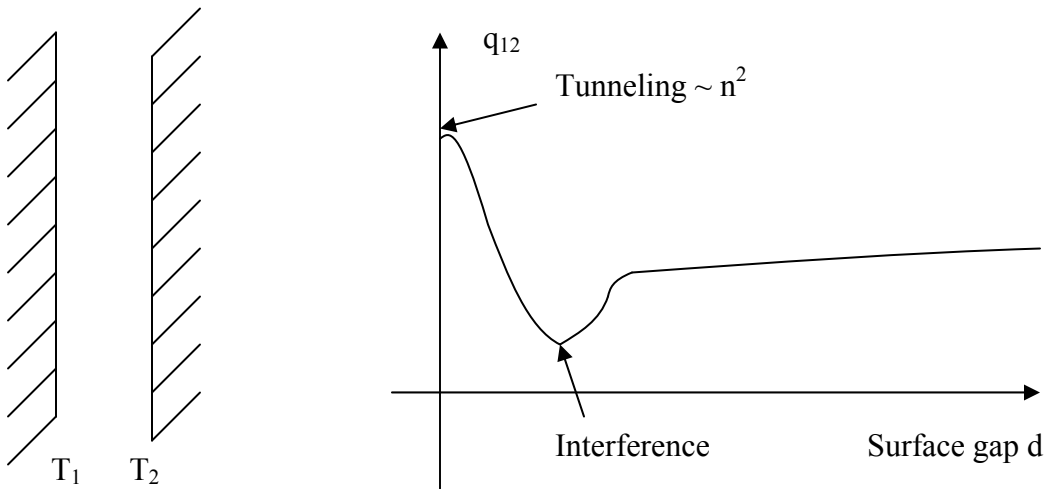
For two blackbody surfaces separated by a gap d, tunneling will occur when d is small. Beyond that region the energy flux follows the fourth-power law, $q_{12} = \sigma(T_1^4 - T_2^4)$. Due to tunneling, the maximum heat flux can exceed the blackbody radiation in vacuum. If the refractive index is n, we have D(E)~n^3 and v~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 $K_e = \dfrac{2e^2}{h}$ (h is the Planck's constant), while for phonons $K_p = \dfrac{\pi^2 k_B T}{3h}$. It should be noted the conductance is independent of materials. The conductance looks like stairs in this case.
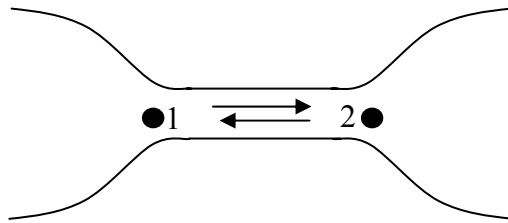
Discussion: What is temperature?
The Landauer formalism is not applicable to special cases where two media are identical and have perfect contact (i.e. merging into a whole bulk material). In these situations, the relationship $q_{12} = K(T_{e1} - T_{e2}) = \dfrac{T_{e1} - T_{e2}}{R_{12}}$ indicates 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.