

1.

(a) The need for Quantum Mechanics arose around the turn of the 20<sup>th</sup> century, spurred on by a number of experimental irregularities. For instance, on the basis of the understanding of nature at the time, the very existence of atoms was puzzling. It was known that atoms consisted of negatively charged electrons "orbiting" positively charged nuclei. From conventional electromagnetism, we know that any accelerating charged particle will emit electromagnetic radiation, and hence lose energy. Therefore, if an electron orbits a nucleus, by the very fact that it is moving around the nucleus, it must be undergoing acceleration. One would expect then, that it would radiate, lose energy and spiral into the nucleus. This clearly doesn't happen, so a theory was needed to explain why not.

Another problem was the phenomenon of photoemission. This is the effect whereby electromagnetic radiation incident on a metal surface causes electrons to be ejected from it. It was observed that electrons are only ejected once the frequency of the light reaches some threshold value, their energy scales linearly with increasing frequency, and the number of electrons ejected increases as the intensity of light increases. This is at odds with classical electromagnetism which predicts that (i) the frequency should not make any difference whatsoever and (ii) the energy of the ejected electrons should only depend on the intensity of light (Energy per unit volume of light is  $\frac{1}{2}\epsilon_0 E^2$ , and Intensity is energy per unit area). There is no reason classically for the number of electrons ejected to depend on intensity. The explanation for this effect within the framework of Quantum mechanics is that light consists not only of waves, but also of particles called *photons*, which contain distinct amounts of energy, depending linearly on the frequency. The proportionality constant is Planck's constant,  $h$ . Electrons are bound to the metal by an energy called the work function, and until the frequency of light is high enough, individual photons don't have enough energy to free them. Above this threshold frequency then, electrons are emitted, and the excess energy is taken up as the electron's kinetic energy. The number of electrons ejected then scales as the number of photons, which depends on the intensity.

Another turning point was the observation of electron diffraction – clear evidence that particles have a wave-like character.

Other experimental evidence was for instance the inability of Physics to explain (i) why some materials are conductors and others insulators; (ii) the origin of electrical resistance; (iii) the observed dependence of specific heat on temperature and (iv) the origin of discrete spectra from heated materials, e.g. the sun.

(b) Wave-packets are used to represent particles, eg. Electrons.

To construct a wave-packet, add two sinewaves:

$$E = E_0 \{ \cos(\omega_1 t - k_1 x) + \cos(\omega_2 t - k_2 x) \}$$

This can be re-written as  $E = 2E_0 \sin(\omega_+ t - k_+ x) \cos(\omega_- t - k_- x)$

Where  $\omega_{\pm} = (\omega_1 \pm \omega_2)/2$  and  $k_{\pm} = (k_1 \pm k_2)/2$

This represents a sinewave of frequency  $\omega_-$  which has its amplitude modulated at the frequency  $\omega_+$ , to produce beats. This therefore exhibits localised areas of enhanced field strength which may be used to represent particles. These localised areas (which we call wave-packets) propagate through space at the group velocity  $\delta\omega/\delta k = c$  (the speed of light). In the absence of dispersion, these wave-packets will propagate without changing shape, but if there is dispersion, they will spread out.

The phase velocity is  $\omega/k = \omega_2/k_2 = c$ . The group velocity can in principle vary, but up to a maximum limit of  $c$ . We could refine the wave-packet by summing together an infinite number of sinewaves, with for eg a Gaussian spectral function. This will increase the localisation.

(c) Construct a wave-packet as follows generally:

$$E = E_0 \int_{-\infty}^{\infty} f(k) e^{i(kx - \omega t)} dk \quad (1)$$

where  $f(k)$  is the spectral function, which is given here as being a gaussian of width  $\delta$ :

$$f(k) = \frac{1}{\sqrt{2\pi\delta}} e^{-\frac{k^2}{2\delta^2}}$$

Therefore, we can write (1) as 
$$E = \frac{E_0}{\sqrt{2\pi\delta}} \int_{-\infty}^{\infty} e^{-\frac{k^2}{2\delta^2} + ik(x-ct)} dk$$

Where we have used the relation that  $\omega = ck$ .

This is known as a Gaussian Integral, and the easiest way to solve is to change the variables to be more manageable: let  $x-ct \rightarrow x'$ ,  $k - ib/2a \rightarrow k'$ , where  $a = 1/2\delta^2$  and  $b = x-ct$ . This gives us

$$E = \frac{E_0}{\sqrt{2\pi\delta}} e^{\frac{-b^2}{4a}} \int_{-\infty}^{\infty} e^{-ak'^2} dk'$$

The integral is just equal to  $\sqrt{\pi/a}$

The net result then is that 
$$E = E_0 \sqrt{\delta} e^{\frac{\delta^2}{2}(x-ct)^2}$$

For a matter wave-packet, the propagation is different due to the fact that unlike em waves, matter waves of different  $k$  have a different momentum,  $p$ , and so travel at

3

different speeds. This leads inevitably to spreading of the wave-packet over time, known as dispersion.

(d) Heisenberg's uncertainty principle states that  $\Delta x \Delta k \geq \frac{\hbar}{2}$ . To see where this comes from, consider the following:

The spectral function has a width  $\Delta k = \delta$ . By inspection, the solution for  $E(x)$  has a width  $\Delta x = 1/\delta$ . Therefore,  $\Delta x \Delta k = 1$ . Given that momentum,  $p = \hbar k$ , it follows that  $\Delta x \Delta p = \hbar$ . This is just a lower bound on the uncertainty, hence the inequality in Heisenberg's uncertainty principle. What this is saying is that we can never simultaneously know absolutely both the position and momentum of a particle, and the more accurately we know one, the less we can know about the other.

2.

(a) Schrödinger's equation can be written in the regions to the left and right of the step as

$$(-\hbar^2/2m\partial^2/\partial x^2) \Psi_I(x) = E\Psi_I(x) \quad \text{Region I}$$

$$(-\hbar^2/2m\partial^2/\partial x^2 + V) \Psi_{II}(x) = E\Psi_{II}(x) \quad \text{Region II}$$

The solutions to these equations are:

$$\Psi_I(x) = A_1 e^{ik_1 x} + B_1 e^{-ik_1 x} \text{ where } k_1 = \frac{\sqrt{2m(E-V)}}{\hbar}$$

$$\text{and } \Psi_{II}(x) = A_2 e^{ik_2 x} \text{ where } k_2 = \frac{\sqrt{2mE}}{\hbar}$$

Matching the wave-functions and their first derivatives at the boundary ( $x = 0$ ) yields the following relationships:

$$A_1 + B_1 = A_2$$

$$ik_1 A_1 - ik_1 B_1 = ik_2 A_2$$

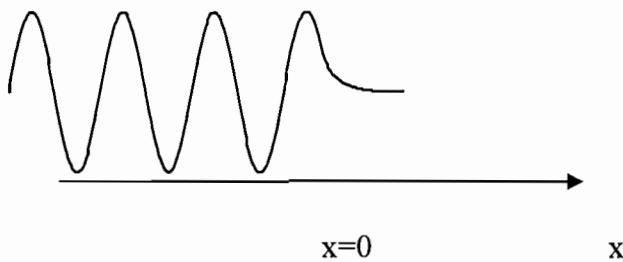
$$\text{i.e. } B_1/A_1 = (k_1 - k_2)/(k_1 + k_2)$$

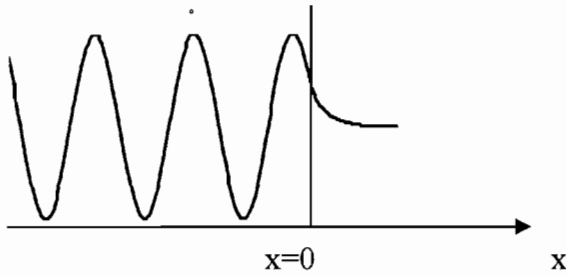
The reflection probability is  $|B_1/A_1|^2$ . This represents the probability that a given particle will be reflected from the potential step, so must lie between 0 & 1. R does *not* represent the fraction of a particle which will be reflected. If we pass a large number of electrons over this potential step, then on average, a fraction R of them will be reflected, but any individual electron will either be totally reflected or totally transmitted.

In this case,  $R = 0.146$ .

(b) Probability density is defined as  $|\psi(x,t)|^2$ . This is the probability of finding the particle at position  $x$ , at time  $t$ .

Wave-function and Probability density:





(c)

Schrodinger's equation is now

$$(-\hbar^2/2m\partial^2/\partial x^2 + V) \Psi_I(x) = E\Psi_I(x)$$

$$(-\hbar^2/2m\partial^2/\partial x^2) \Psi_{II}(x) = E\Psi_{II}(x)$$

$$(-\hbar^2/2m\partial^2/\partial x^2 + V) \Psi_{III}(x) = E\Psi_{III}(x)$$

**Region I****Region II****Region III**

The solutions to these equations are:

$$\Psi_I(x) = A_1 e^{ik_1 x} + B_1 e^{-ik_1 x} \text{ where } k_1 = \frac{\sqrt{2m(E-V)}}{\hbar}$$

$$\Psi_{II}(x) = A_2 e^{ik_2 x} + B_2 e^{-ik_2 x} \text{ where } k_2 = \frac{\sqrt{2mE}}{\hbar}$$

And  $\Psi_{III}(x) = A_3 e^{ik_1 x}$  where  $k_3 = k_1$ 

Matching the wave-functions and their first derivatives at the boundary ( $x = 0$ ) yields the following relationships:

$$A_1 + B_1 = A_2 + B_2$$

$$ik_1 A_1 - ik_1 B_1 = ik_2 A_2 - ik_2 B_2$$

matching at  $x = L$  gives:

$$A_2 e^{ik_2 L} + B_2 e^{-ik_2 L} = A_3 e^{ik_1 L}$$

$$ik_2 A_2 e^{ik_2 L} - ik_2 B_2 e^{-ik_2 L} = ik_1 A_3 e^{ik_1 L}$$

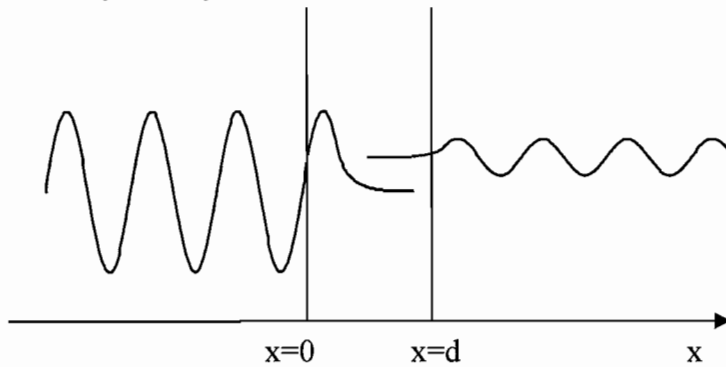
combining these equations, we eventually obtain for R:

$$R = |B_1/A_1|^2 = \left[ \frac{[(k_1^2 - k_2^2)(1 - e^{2ik_2 L}]}{[(k_2 + k_1)^2 - (k_2 - k_1)^2 e^{2ik_2 L}]}\right]^2$$

$$= \left[ 1 + \frac{4k_1^2 k_2^2}{(k_1^2 - k_2^2) \sin^2 k_2 L} \right]^{-1}$$

= 0.357

Probability density:



If we say that the amplitude of the incident wave is 1, then the wave-function in the barrier region is:

Assumptions: (i) only an exponentially decaying term in barrier, (ii) effective mass of electron in each region is the same

To improve precision, would include full form of  $\psi_{II}$ , and would determine transmission probability for entire structure.

(d) For a wave-packet, there are a range of energies depending on the initial spatial extent of the packet. A packet which is narrow will be made up of plane waves with a wide range of momenta, and so with a wide range of plane wave energies, so the transmission and reflection probabilities will be different for the packet than the single plane wave.

3.

There is another way to achieve very similar behavior: if a conventional p-n junction diode is doped heavily enough ( $\sim 10^{25}$  dopants  $m^{-3}$ ), it is possible to cause the Fermi levels in the n and p-type materials to be in the conduction and valence bands, respectively, as shown in the Figure below. Also, the effect of very high doping levels is to make the depletion region extremely thin, in the nm range, so appreciable tunnelling can occur.

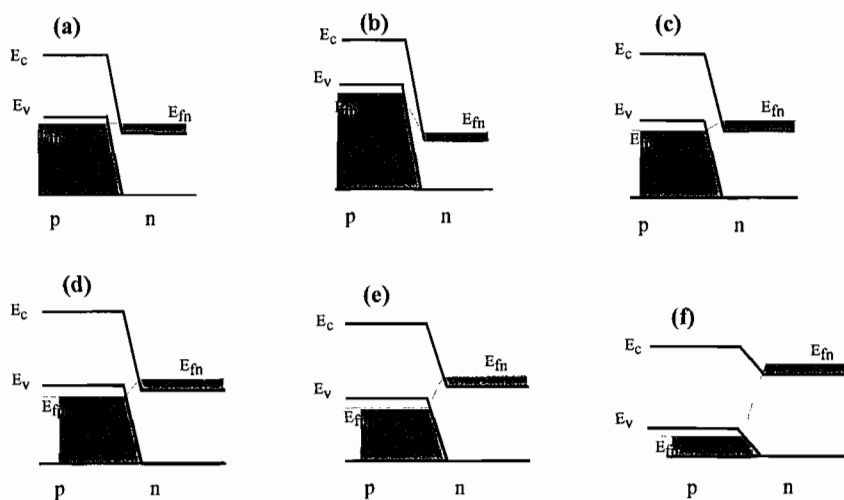


Fig. 2.37. Band diagram of tunnel diode.

In (a), i.e. under zero applied bias: there is no net current flow, as the electron current from the conduction band of the n-type into the valence band of the p-type is balanced by the electron current from the valence band of the p-type in to the conduction band of the n-type. In Fig. (b), under reverse bias conditions, the bands on the p-type side are raised relative to the n-type side, and electrons can flow from p to n, tunnelling across the depletion region. The width of this region will increase as the voltage is increased, so little current will actually flow. In Fig. (c), which is under a low forward bias, the electron-filled states in the n-type conduction band overlap with the holes in the p-type valence band and a significant current can tunnel across the depletion region, similar to Regime B in Fig. (b). In Fig. (d), as the forward bias is increased, the degree of overlap between the n-type conduction band electrons and the p-type valence band holes decreases, as more of them start to overlap with the band gap within the p-type. This has the effect of reducing the current across the depletion region as there are fewer states for the n-type electrons to tunnel into. In Fig. (e), similar to Regime (c) in Fig. (b), the current drops to its minimum value, as there is no longer any overlap between the conduction band electrons in the n-type and holes in the p-type: there are no available states for the electrons to tunnel into. The only current which can flow at this point is a small inelastic tunnel current and a small thermal diffusion current. In Fig. (f), when the applied forward bias is large enough, the height of the potential barrier between the n and p-type is low enough for a thermal diffusion current to flow over the barrier, and this becomes the dominant means of current flow. Resonant tunnelling diodes initially gained a lot in interest for their potential application in oscillator circuits, particularly ones operating at high (Microwave) frequencies. The reason for this can be seen by considering the simplest possible oscillator: an LC circuit (i.e. an inductor in parallel with a capacitor). Due to the phase difference of 180 degrees between the voltage dropped across each of these, energy is effectively continually transferred from one component to the other — the circuit is an oscillator. Once the oscillations begin, if we remove the voltage driving source, the oscillations would continue indefinitely in the absence of any resistance within the circuit. However, all circuits have some resistance, so real oscillator circuits have a finite Q-factor. In principle, if we could add a negative resistance into the circuit to counteract the stray resistance of the components, we could greatly increase the circuit's Q-factor. This is done by adding a resonant tunnelling diode into the LC circuit, and ensuring that it is operating in the middle of its NDR region. This is illustrated in the Figure below. In recent years, the tunnel diode has been replaced by digital components which are more reliable and which have significantly better performance.

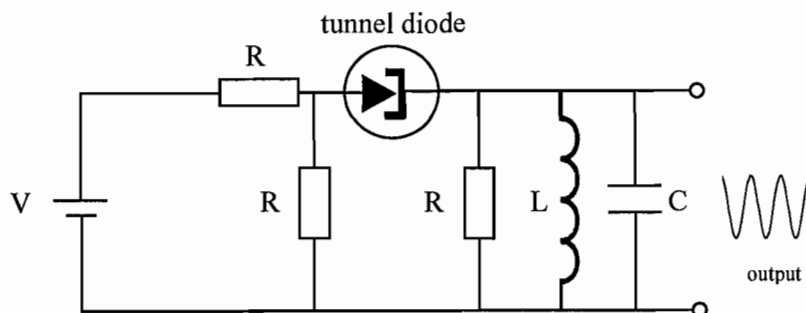


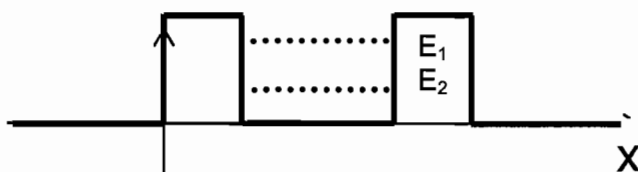
Fig. 2.38. Typical circuit utilising a tunnel diode. The voltage source  $V$  is used to set the diode operating in the NDR region (between A and B in Fig. 2.33(b)), and to start the oscillation. It also provides the energy to sustain the oscillation of the circuit. The oscillation frequency is

$$\left( \frac{1}{2\pi\sqrt{LC}} \right)$$

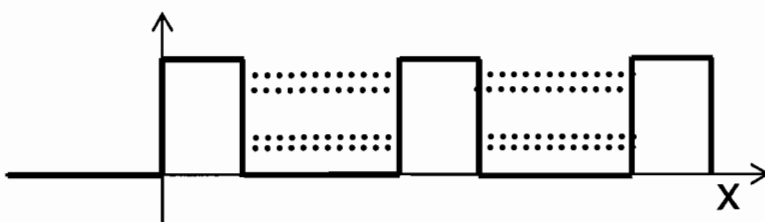


4.

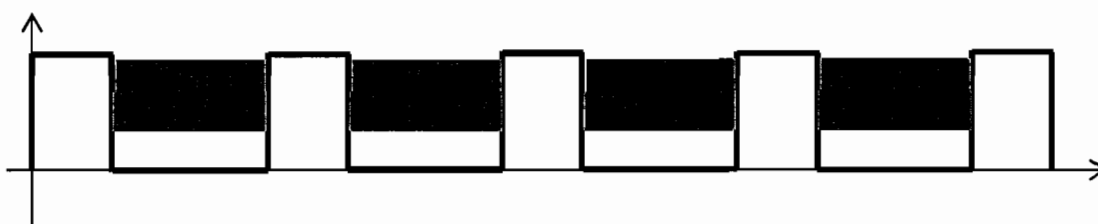
(a)



Now add in another well:



Now for a very large number of wells:



Energy levels become bands.

If the wells are close enough, the tail of the wave-functions within each well can extend appreciably into neighbouring wells. This gives rise to a coupling between the wells, and generates new wave-functions which are combinations of the original ones. For two wells, those combinations are the sum and difference between the original functions. Hence, there will be two states instead of one. For  $N$  wells, each state will become  $N$  closely spaced states. The closer the wells are to each other, the larger will be the splitting in energy. This is similar to beats in the addition of waves: the closer the frequency, the more pronounced are the beats. As the wells are brought closer, the coupling increases and so therefore does the splitting of the energy levels.

(b) Assumptions in the nearly-free electron model:

The electrons do not interact with each other

The electrons have *discrete* energy values

The electron gas follows Fermi-Dirac statistics

The general solution of the Schrödinger equation with a periodic potential is  $\psi(x) = e^{ikx}u(x)$ . This is a plane wave modulated by the function  $u(x)$ , where  $u(x)$  is a periodic function with the periodicity of the lattice, i.e.  $u(x)$  represents the influence of the crystal potential. This is known as **Bloch's theorem**, and  $u(x)$  as a **Bloch function**.

If we expand the potential as a Fourier series, we can now do the same for  $u(x)$ , to obtain:

$$u(x) = \sum_n C_n e^{iG_n x}$$

where  $n = 0, \pm 1, \pm 2, \dots$  and  $G_n = 2\pi n/a$

That gives for the total expansion of the wave function:

We now insert the Fourier expansions of both  $\psi(x)$  and  $u(x)$  into Schrodinger's equation,

$$\psi(x) = \sum_n C_n e^{i(k+G_n)x}$$

$$(-\hbar^2/2md^2/dx^2 + V) \Psi(x) = E\Psi(x)$$

We end up with a set of simultaneous equations in the unknown  $C_n$ . Note that the  $V_p$  are known, as the form of the crystal potential is assumed initially. There are an infinite number of terms, so to make the problem manageable, we artificially truncate the series and consider only the leading-order terms given by  $n = 0, \pm 1$ . This is justified for weak potentials such as those found in metals.

$$V(x) = V_0 + V_1 e^{iG_1 x} + V_{-1} e^{iG_{-1} x}$$

If we continue along the same lines, we can assume that the wave-function also only contains leading-order terms, i.e.

$$\psi(x) = [C_0 + C_1 e^{iG_1 x} + C_{-1} e^{iG_{-1} x}] e^{ikx}$$

$$(-\hbar^2/2md^2/dx^2 + V_0 + V_1 e^{iG_1 x} + V_{-1} e^{iG_{-1} x}) [C_0 + C_1 e^{iG_1 x} + C_{-1} e^{iG_{-1} x}] e^{ikx} =$$

$$E [C_0 + C_1 e^{iG_1 x} + C_{-1} e^{iG_{-1} x}] e^{ikx}$$

If we just consider a region where  $C_0$  and  $C_{-1}$  dominate, we are left with the relationships (noting that  $G_{-1} = -G_1$  etc.):

$$(-\hbar^2 k^2 C_0 / 2m + V_0 C_0 + C_0 V_1 e^{iG_1 x} + C_0 V_{-1} e^{iG_{-1} x} - \hbar^2 (k + G_{-1})^2 C_{-1} e^{iG_{-1} x} / 2m +$$

$$V_0 C_1 e^{iG_{-1}x} + V_1 C_1 + V_{-1} C_{-1} e^{2iG_{-1}x} = EC_0 + EC_{-1} e^{iG_{-1}x}$$

Collecting terms in  $e^{iG_{-1}x}$ , we find that:

$$C_0 V_{-1} = [\hbar^2(k + G_{-1})^2/2m + E - V_0] C_{-1}$$

Terms without any exponent give:

$$C_{-1} V_1 = [(\hbar^2 k^2/2m) + E - V_0] C_0$$

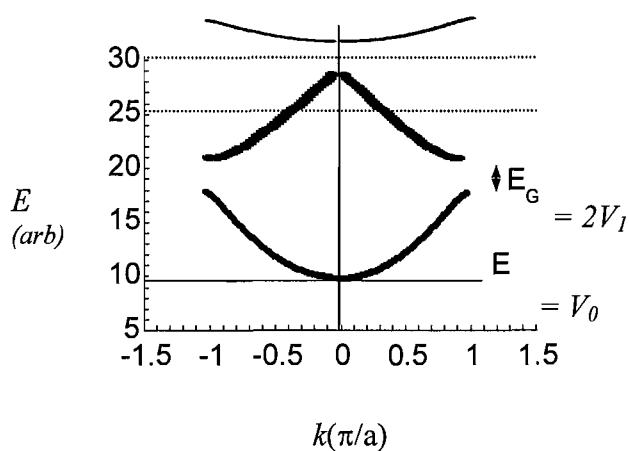
For a non-trivial solution, both ratios for  $C_{-1}/C_0$  must be equal, i.e.

$$\begin{aligned} C_{-1}/C_0 &= [(\hbar^2 k^2/2m) + E - V_0]/V_1 \\ &= V_{-1}/[\hbar^2(k + G_{-1})^2/2m + E - V_0] \end{aligned}$$

or,  $[E - V_0 + (\hbar^2 k^2/2m)] [E - V_0 + \hbar^2(k + G_{-1})^2/2m]$

$$= V_1 V_{-1} = |V_1|^2 \quad (\text{Everything is symmetric})$$

(c) Dispersion relation:



The wave-functions at the valence and conduction-band edges can be written as:

12

$$\psi_v(x) = 2C_0 \cos\left(\frac{\pi x}{a}\right) \quad \text{and} \quad \psi_c(x) = 2iC_0 \sin\left(\frac{\pi x}{a}\right)$$

These are standing waves which means electrons at these energies do not travel through the crystal, and hence do not contribute to conduction.

5. Moore's law states that the number of transistors in a microprocessor is doubling every 16-18 months.

(a) Answer should include a discussion of several of the following topics:

- Better control of fabrication techniques and materials
- Vacuum devices to solid-state
- Increase in electron/hole mobility
- Reduction in size of transistors
- Band engineering, and novel device structures

The desire to have smaller transistors has two reasons: (i) smaller means faster, as electrons have shorter distances to travel, and there is less scattering, and (ii) smaller means higher density, so high performance devices can be made portable.

(b) **Quantum effect:** Tunnelling. Explain what tunnelling is. There is a finite probability that a particle of energy  $E$  striking a barrier of height  $V$  where  $V > E$ , will get through! This is non-classical behaviour. Classically, we expect that at the point where  $E = V$ , the particle will have zero kinetic energy, and will reverse its trajectory. Classically, you cannot have a situation where the particle has negative net kinetic energy, so the particle should not be able to penetrate the barrier. This purely quantum effect is called *tunnelling*, and is responsible for a number of effects, such as nuclear  $\alpha$ -decay and field emission.

Description should include some schematic of wave-function decay into "forbidden" regions. Then, tunnelling is a problem in transistors because gate oxides are becoming thinner, and are at the nm level. To continue reducing dimensions will lead to more tunnelling. This can be overcome by the use of high- $k$  dielectrics.

**Classical effect:** Electromigration. This is an effect whereby current flow causes atoms in wires to move to such an extent that the wires eventually fail. This affects all current-carrying wires. This limits the lifetime of interconnects in ICs, and as transistors and hence interconnects shrink, the lifetime will decrease further. This can be overcome by using materials which have low surface diffusion constants, and by coating the wires with a passivation layer.

(c) Reason: transistors cannot continue to shrink and get faster indefinitely, and as they get smaller, the reproducibility in properties will decrease due to the statistical spread in doping levels.

(i) Resonant tunnelling devices use band engineering to produce double-barrier structures a few nm apart, and rely on atomic-level manufacturing precision (MBE). Therefore, they are

extremely difficult to make in bulk quantities, and are not very reproducible. They are however, very fast, and because they exhibit NDR, they are useful as microwave sources.

(ii) hot-electrons are electrons which have excess kinetic energy, which they gain from passing from a region of high potential to a region of low potential. This essentially means a heterojunction, which is easy to fabricate. Unfortunately, these hot electrons lose their excess energy very quickly after scattering, so to utilise them in devices, the gate length must be shorter than the mean-free path. As they are travelling ballistically (at the Fermi velocity which is at least an order of magnitude faster than the drift velocity) devices using hot electrons are extremely fast.

(iii) In molecular electronic devices, the functional elements are molecules, which will be identical, so they should all have the same properties. As molecules are so small, they can conduct electric current extremely quickly. Also, they exhibit a vast range of transport phenomena. Answer should perhaps include a schematic of what such a device might look like.