# ENGINEERING TRIPOS PART IIA 2004

Solutions to Module 3B5
Semiconductor Engineering
Principal Assessor: Professor P Migliorato
Second Assessor: Dr A J Flewitt

Q.1

a) $\psi(x) = A\sin kx + B\cos kx$

$$k^2 = \frac{2mE}{\hbar^2}$$

$\psi(a) = 0 \Longrightarrow A\sin ka = 0$

$ka = n\pi \qquad n = 1, 2, \ldots$

$$E_n = n^2\frac{\hbar^2\pi^2}{2ma^2}$$

$\psi(x) = A\sin n\frac{\pi}{a}x$

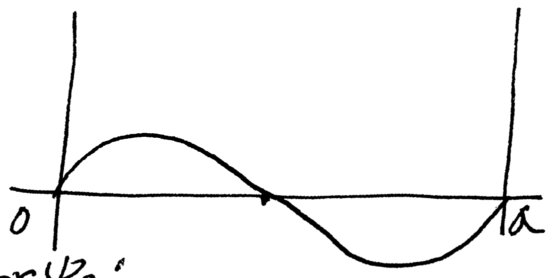
Determine A from normalization:

$$\int_0^a |\psi|^2 dx = |A|^2\int_0^a \sin^2 n\frac{\pi}{a}x\, dx = 1$$

$$A = \sqrt{\frac{2}{a}}$$

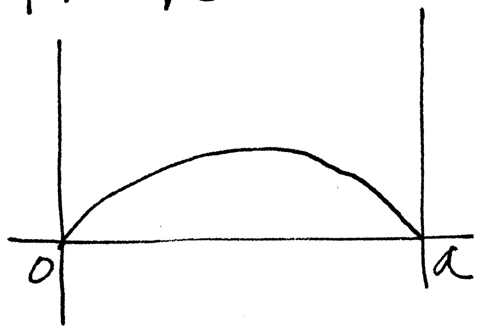
b) For $n = 2$

$\psi_2(x) = \sqrt{\frac{2}{a}} \sin \frac{2\pi}{a} x$



For $n = 1$

$\psi_1 = \sqrt{\frac{2}{a}} \sin \frac{\pi}{a} x$



For $\psi_2$:

c) $\int_0^{a/4} |\psi_2|^2 dx = \frac{2}{a} \int_0^{a/4} \sin^2 \frac{2\pi}{a} x \, dx$ etc.

long handed $= \frac{1}{4}$

or by inspection!

same for probability between $\frac{3a}{4}$ and $a$

For $\psi_1$:

$\int_0^{a/4} |\psi_1|^2 dx = \frac{1}{4} - \frac{1}{2\pi} < \frac{1}{4}$

By inspection $\int_{3/4 a}^{a} |\psi_1|^2 dx$ is the same

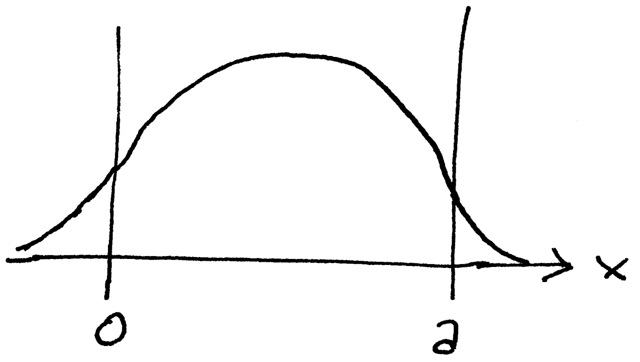
d) For $0 < x < a$

$\psi(x) = A \sin kx + B \cos kx$

For $-\infty < x < 0$ $\quad \psi(x) = C e^{\alpha x}$

For $a < x < \infty$ $\quad \psi(x) = D e^{-\alpha x}$

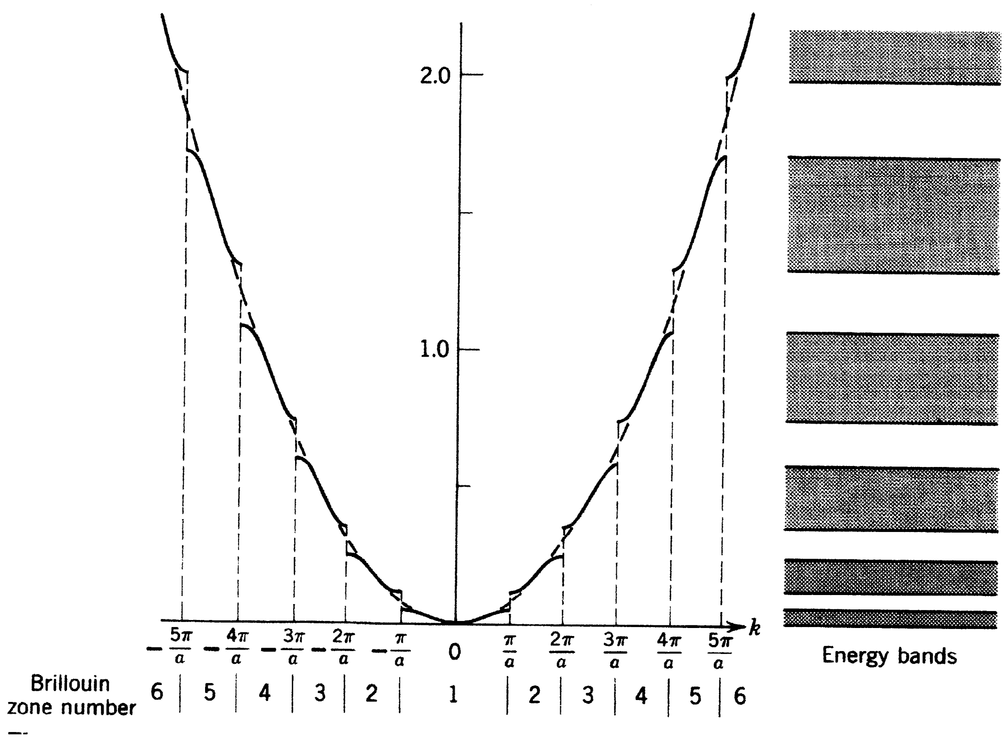
where: $\alpha = \frac{\sqrt{2m(V_0 - E)}}{\hbar}$



In this case (finite depth of the well) one cannot assume that $\psi(0) = \psi(a) = 0$, hence the wavefunction extends outside the potential well. This is a typical quantum-mechanical effect: 'tunnelling'.

Q.2

a)



When Bragg's condition is verified incident and reflected wave form standing waves. There are two types of standing waves:

$$e^{ikx}+e^{-ikx} \rightarrow \cos kx=\cos n\frac{\pi}{d}a$$

$$e^{ikx}-e^{-ikx} \rightarrow \sin kx=\sin n\frac{\pi}{d}a$$

The two standing waves differ substantially in the location of the nodes and therefore in the maxima and minima of the probability density $|\psi^2|$. The cosine wave has maxima for x=ma, m=1,2,3…, that is at the crystal ions positions. The sine wave has nodes for x=ma.

So an electron which is in a state represented by the cosine wave, experiences a *different crystal potential* than an electron in a state represented by the sine wave. This is tantamount to saying that the eigenvalues corresponding to the two eigenfunctions must be different. This is the origin of the energy gaps in the nearly-free electron approximation b)

$$v = \frac{p}{m} = \frac{\hbar k}{m} \tag{1}$$

$$E = \frac{p^2}{2m} + V(x) = \frac{\hbar^2 k^2}{2m} + V(x) \tag{2}$$

From Equations (8) and (9):

$$v = \frac{1}{\hbar}\frac{dE}{dk} \tag{3}$$

since V(x) is independent of k. Equation (3) shows that the electron velocity depends on the actual E-k curve.

The classical part expresses dE as the work done by a classical particle travelling a distance **vdt** under the influence of a force **eF**, yielding:

$$dE = eFvdt = eF\frac{1}{\hbar}\frac{dE}{dk}dt \tag{4}$$

We obtain the acceleration by differentiating Equation 3 as follows

$$\frac{dv}{dt} = \frac{1}{\hbar}\frac{d}{dt}\left(\frac{dE}{dk}\right) = \frac{1}{\hbar}\frac{d^2E}{dk^2}\frac{dk}{dt} \tag{5}$$

Expressing now dk/dt from Equation 4

$$\frac{dk}{dt} = \frac{1}{\hbar}eF$$

and substituting it into Equation (5) we get

$$\frac{dv}{dt} = \frac{1}{\hbar^2}\frac{d^2E}{dk^2}eF \qquad (6)$$

Comparing this formula with that for a free, classical particle

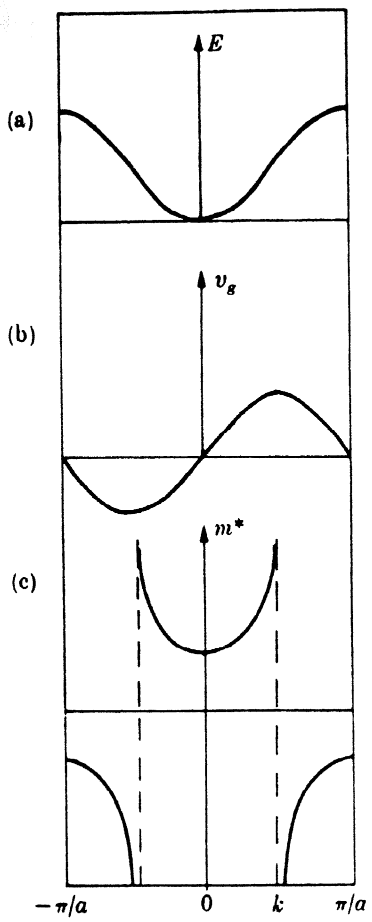
$$m\frac{dv}{dt} = eF \qquad (7)$$

we define

$$m^* = \hbar^2\left(\frac{d^2E}{dk^2}\right)^{-1} \qquad (8)$$

as the effective mass of an electron.

c)



*Electron energy, velocity and effective mass as a function of k*

The surprising result that the effective mass is negative for k approaching $\pm\frac{\pi}{a}$ can be understood if we consider what happens when an electron initially with $k < \frac{\pi}{a}$ is subjet to an electric field producing a force in the direction of positive k: the electron will be accelerated and its momentum, hence the wavevector k, will increase. The point representing the particle in the E-k diagram will move towards $\frac{\pi}{a}$. As long as the electron is far away from $\frac{\pi}{a}$, it is represented by the wavefunction of a free particle:

$$\psi_+ = e^{ikx}$$

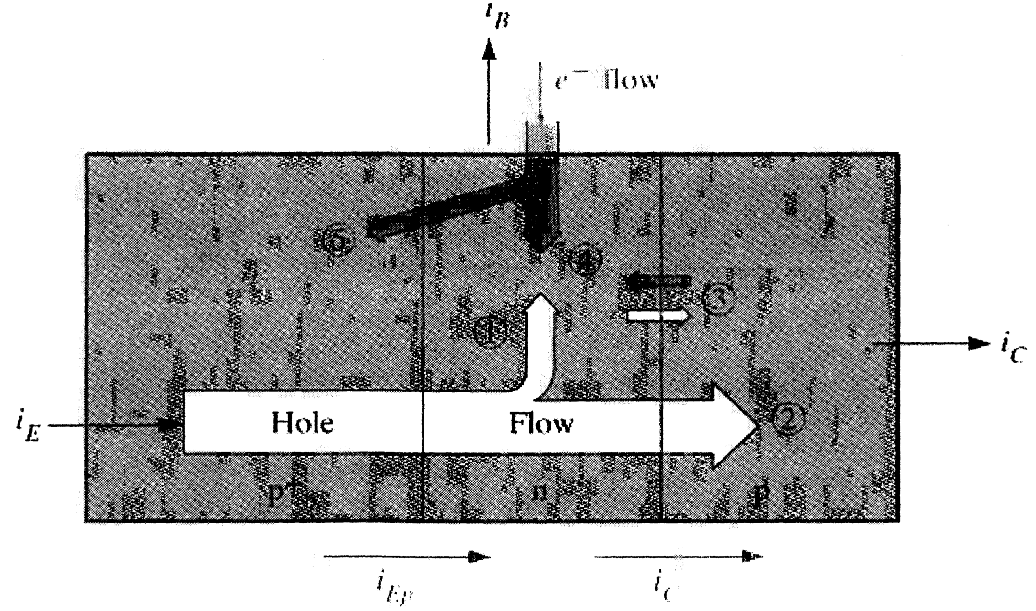
But, we know what happens for $k = \frac{\pi}{a}$: Bragg diffraction will take place, the electron will be reflected back and its wavefunction now becomes a standing wave:

$$\psi = \psi_+ + \psi_- = e^{ikx} + e^{-ikx}$$

Since $|\psi_-|^2$ represents the probability density of the state with momentum -k (the reflected wave), this result can be viewed as if the electron changed its momentum from k to -k, *while a force in the direction of positive k is applied.* This is only possible if the electron has at that point a negative mass.

3B5 - 2004

Q.3 (a) The p$^+$np BJT essentially consists of two pn junction back to back. The current flow under normal operation (in which the emitter and collector are biased positively and negatively respectively with respect to the base) is shown below.



Let us first consider the junction between the emitter and the base. This junction is under forward bias. Therefore, a large hole current flows from the p$^+$ emitter into the n base. As the base is less heavily doped than the emitter, a smaller electron current (5) flows from the base into the emitter (ideally this current should be as small as possible for a high gain to be achieved). The collector-base junction, on the other hand, is under reverse bias, and so a small reverse bias electron and hole current flows across this junction (3). However, if the width of the base is small compared with the diffusion length of electrons in the n-type semiconductor, then a significant number of holes injected into the base from the forward biased emitter-base junction will diffuse across the base to the collector junction. These holes will be swept across the depletion region between the base and the collector to yield a large hole current (2). A proportion of the holes injected into the base from the emitter will recombine with electrons in the base (1), and so an electron current must flow into the base to compensate for the recombined electrons (4). If an insufficient electron current (4) is supplied, then the base would gain a net positive space charge, causing a reduction in the injected hole current. Therefore, as a constant proportion of holes will recombine in transit across the base, it is this small base electron current (4) which limits the large hole current flowing into the collector (2).

(b) (i) The Continuity equation for holes is

$$\frac{\partial(\Delta p)}{\partial t} = -\frac{\Delta p}{\tau_h} - \mu_h \varepsilon \frac{\partial(\Delta p)}{\partial x} + D_h \frac{\partial^2(\Delta p)}{\partial x^2}$$

where $\Delta p$ is the number density of excess holes, $\tau_h$ is the hole lifetime, $\mu_h$ is the hole mobility, $\varepsilon$ is the electric field and $D_h$ is the hole diffusion coefficient. We assume that there is no electric field outside the depletion regions and that the system is in a steady state (there is no time dependence), and so in the base this equation reduces to

$$\frac{\Delta p_n}{\tau_h} = D_h \frac{\partial^2(\Delta p_n)}{\partial x^2}$$

which has the general solution,

$$\Delta p_n(x) = A\exp\left(\frac{-x}{L_h}\right) + B\exp\left(\frac{x}{L_h}\right)$$

where $A$ and $B$ are constants and $L_h = (D_h\tau_h)^{1/2}$ is the diffusion length of holes. The constants can be determined if we assume that the excess hole concentration at the junction with the collector is zero ($\Delta p_n(W_b) = 0$). Differentiating the general solution twice gives

$$\frac{\partial^2(\Delta p_n)}{\partial x^2} = \frac{A}{L_h^2}\exp\left(\frac{-x}{L_h}\right) + \frac{B}{L_h^2}\exp\left(\frac{x}{L_h}\right)$$

Hence, at $x = 0$,

$$\frac{\Delta p_n(0)}{\tau_h} = \frac{D_h}{L_h^2}(A+B)$$

and as $L_h = (D_h\tau_h)^{1/2}$ then

$$\Delta p_n(0) = A + B$$

At $x = W_b$,

$$0 = A\exp\left(\frac{-W_b}{L_h}\right) + B\exp\left(\frac{W_b}{L_h}\right)$$

$$B\exp\left(\frac{W_b}{L_h}\right) = -A\exp\left(\frac{-W_b}{L_h}\right)$$

$$(\Delta p_n(0) - A)\exp\left(\frac{W_b}{L_h}\right) = -A\exp\left(\frac{-W_b}{L_h}\right)$$

$$1 - \frac{\Delta p_n(0)}{A} = \exp\left(\frac{-2W_b}{L_h}\right)$$

$$\frac{\Delta p_n(0)}{A} = 1 - \exp\left(\frac{-2W_b}{L_h}\right)$$

$$A = \frac{\Delta p_n(0)}{1 - \exp\left(\frac{-2W_b}{L_h}\right)}$$

and hence

$$B = \Delta p_n(0) - \frac{\Delta p_n(0)}{1 - \exp\left(\frac{-2W_b}{L_h}\right)}$$

$$B = \Delta p_n(0)\left(\frac{1 - \exp\left(\frac{-2W_b}{L_h}\right) - 1}{1 - \exp\left(\frac{-2W_b}{L_h}\right)}\right)$$

$$B = \Delta p_n(0)\left(\frac{-\exp\left(\frac{-2W_b}{L_h}\right)}{1 - \exp\left(\frac{-2W_b}{L_h}\right)}\right)$$

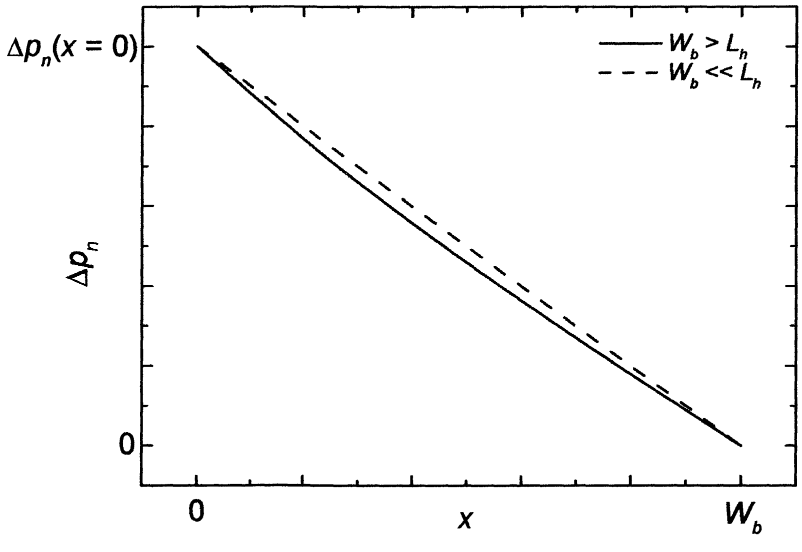
Therefore,

$$\Delta p_n(x) = \frac{\Delta p_n(0)}{1 - \exp\left(\frac{-2W_b}{L_h}\right)}\exp\left(\frac{-x}{L_h}\right) + \Delta p_n(0)\left(\frac{-\exp\left(\frac{-2W_b}{L_h}\right)}{1 - \exp\left(\frac{-2W_b}{L_h}\right)}\right)\exp\left(\frac{x}{L_h}\right)$$

$$\Delta p_n(x) = \Delta p_n(0)\left\{\frac{\exp\left(\frac{-x}{L_h}\right) - \exp\left(\frac{-2W_b}{L_h}\right)\exp\left(\frac{x}{L_h}\right)}{1 - \exp\left(\frac{-2W_b}{L_h}\right)}\right\}$$

$$\underline{\Delta p_n(x) = \Delta p_n(0)\left\{\frac{\exp\left(\frac{W_b - x}{L_h}\right) - \exp\left(\frac{x - W_b}{L_h}\right)}{\exp\left(\frac{W_b}{L_h}\right) - \exp\left(\frac{-W_b}{L_h}\right)}\right\}}$$

(ii) The excess hole concentration is shown below



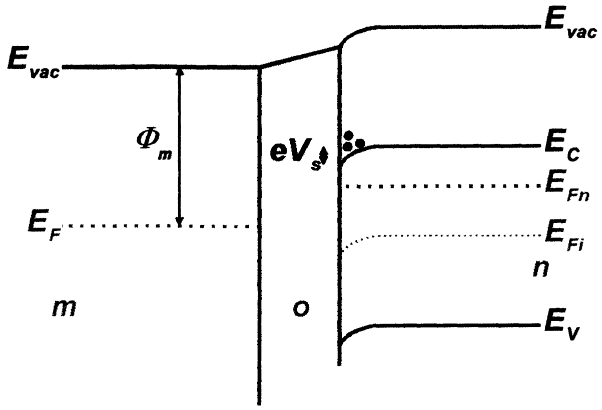
If $W_b << L_h$, then there will be almost no recombination in the base and the current flow (the gradient of the above line) will be constant, and the excess hole concentration will drop linearly across the base. If $W_b > L_h$, then there will be significant recombination of carriers in the base, and the current will fall between the emitter and the collector. The excess carrier concentration then follows a curve below the straight line. For a high gain device, recombination should be minimised, so $W_b << L_h$ is the preferable situation.

Q4 ~~2~~ (a) (i) The flat band condition is when there is no bending of the conduction or valence bands in the semiconductor. For a perfect, ideal MOS structure, this occurs when there is no potential difference between the gate and channel. However, if the workfunctions of the metal gate and semiconducting channel are different, then the voltage required on the gate to achieve the flat band condition is

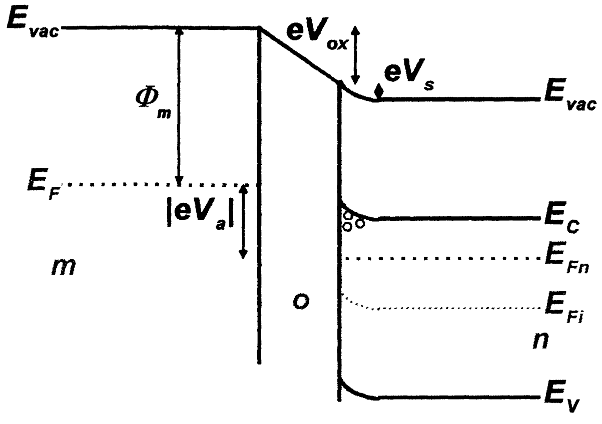
$$V_{FB} = \frac{\phi_m - \phi_{sc}}{e}$$

Any other trapped, fixed or mobile charges in the oxide layer will also affect the voltage required to achieve the flat band condition.

(ii) Accumulation occurs when a voltage is applied to the gate such that extra majority carriers are attracted to the channel layer in the semiconductor below the insulator. In the case of the NMOS device, a positive gate voltage is required to achieve this, as shown:

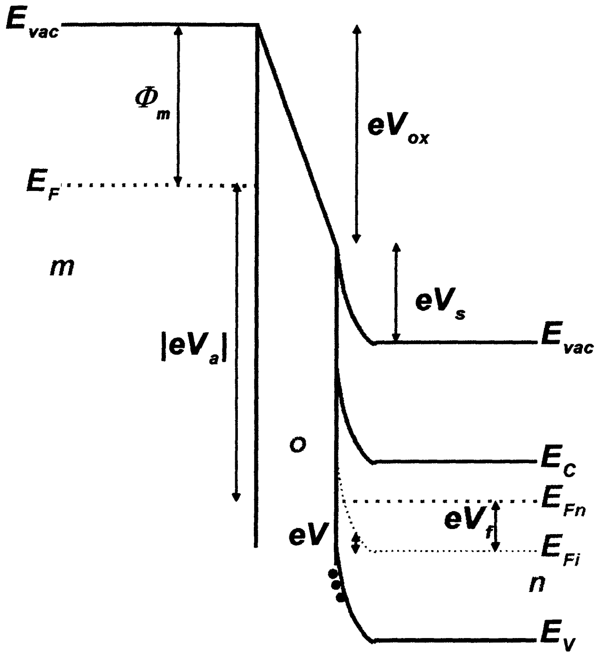


(iii) Depletion occurs when a voltage is applied to the gate such that majority carriers are repelled away from the channel layer in the semiconductor. In the case of the NMOS structure, a negative gate voltage is required:



(iv) Inversion occurs when the bending of the conduction and valence bands close to the insulator is sufficient that, in the case of the NMOS device, the Fermi level is closer to the valence band than the conduction band, and so there are more minority carriers (holes) in the channel than majority carriers (electrons). A negative gate voltage is required to achieve this in an NMOS device.

(v) If a sufficiently large negative voltage is applied to the gate, then band bending may be sufficient that the Fermi level is as close to the valence band in the channel as it is to the conduction band in the bulk of the semiconductor. There are a large number of minority carriers in the channel:



(b) We know that

$$n = n_i \exp\left(\frac{E_F - E_i}{kT}\right)$$

For a n-doped semiconductor, assuming that all dopants are ionised,

$$N_D = n_i \exp\left(\frac{E_F - E_i}{kT}\right)$$

$$E_F - E_i = kT \ln\left(\frac{N_D}{n_i}\right)$$

The surface potential at strong inversion is then

$$V_s = \frac{2(E_F - E_i)}{e}$$

$$= 2\frac{kT}{e}\ln\left(\frac{N_D}{n_i}\right)$$

$$= 2.0\cdot 862\times10^{-4}.298\ln\left(\frac{10^{22}}{1\cdot5\times10^{16}}\right)$$

$$\underline{V_s = 0\cdot688\ \mathrm{V}}$$

(c) The Gauss Law of Electrostatics states that

$$\mathrm{div}\mathbf{E} = \frac{\rho}{\varepsilon_0 \varepsilon_r}$$

$$\frac{dE}{dx} = \frac{eN_D}{\varepsilon_0 \varepsilon_r}$$

where it has been assumed that all donors are ionised in the depletion region. Integrating,

$$E = \int \frac{eN_D}{\varepsilon_0 \varepsilon_r} \mathrm{d}x$$

$$E = \frac{eN_D x}{\varepsilon_0 \varepsilon_r} + C$$

Let us take $x = 0$ to be at the oxide-semiconductor interface. Assuming that there are no electric fields in the semiconductor outside the depletion region, we can use the boundary condition that $E = 0$ at $x = w$, so

$$E = \frac{eN_D (x - w)}{\varepsilon_0 \varepsilon_r}$$

Now,

$$V = -\int E \mathrm{d}x$$

$$= \frac{eN_D}{\varepsilon_0 \varepsilon_r} \int (w - x) \mathrm{d}x$$

$$V = \frac{eN_D \left( wx - \frac{x^2}{2} \right)}{\varepsilon_0 \varepsilon_r} + D$$

We use the boundary condition that $V = 0$ at $x = w$, so

$$V = \frac{eN_D \left( wx - \frac{x^2}{2} - \frac{w^2}{2} \right)}{\varepsilon_0 \varepsilon_r}$$

Hence, the surface potential at $x = 0$ is

$$V_{si} = \frac{eN_D w^2}{2\varepsilon_0 \varepsilon_r}$$

and so the width of the depletion region at strong inversion is

$$\underline{w=\left(\frac{2\varepsilon_0\varepsilon_r V_{si}}{eN_D}\right)^{1/2}}$$

For the device described previously,

$$\begin{aligned} w &= \left(\frac{2\varepsilon_0\varepsilon_r V_{si}}{eN_D}\right)^{1/2} \\ &= \left(\frac{2.8\cdot 854\times 10^{-12}.11\cdot 8.0\cdot 688}{1\cdot 602\times 10^{-19}.10^{22}}\right)^{1/2} \\ &\underline{w = 300\ \text{nm}} \end{aligned}$$

(d) The width of the depletion region calculated previously applied for the quasi-static case. If the gate voltage is varying with time, then the depletion region may be greater than this. This is known as deep depletion, and occurs because holes are created thermally in the channel, and some time is required for the inversion layer to form. If the gate voltage is varied more quickly than this formation time, then depletion may extend further into the bulk of the semiconductor.