Models for charge transport in photovoltaic devices

Giles Richardson*

January 25, 2016

1 Introduction

The aim of this short course is to give you some idea of the different approaches used to model charge transport in photovoltaic devices. The style of presentation reflects my background in continuum mechanics (*e.g.* fluid mechanics, elasticity etc.) and so is rather different to that encountered in a solid-state physics course. Nevertheless despite the difference in approach the resulting models are the same and I hope that you will find it helpful to see things from a different perspective. It must be emphasised that the list of modelling approaches described here is far from comprehensive.

After the Introduction the material is divided into 4 further sections. In Section 2 we briefly review the equilibrium statistical mechanics of a semiconductor as this has some bearing on the non-equilibrium models that we consider in the subsequent 3 sections. In Section 3 we briefly discuss hopping models of charge transport and relate these to continuum drift-diffusion model of the same process. In Section 4 we elaborate on the drift-diffusion model introduced in the previous section and apply this to the particular example of an n-p homojunction used as a solar cell. Furthermore we show that the current-voltage curve predicted by modelling this device using a drift-diffusion model is accurately approximated by a Shockley equivalent circuit. This leads us, in Section 5, to discuss Shockley equivalent circuit models of solar cells and their use in inferring device physics from current-voltage data.

2 Equilibrium electron and hole distributions in a semiconductor

We begin by briefly discussing the equilibrium distribution of charge carriers in a conventional (inorganic) semiconductor before going on to consider charge transport processes. The band diagram of a semiconductor (*i.e.* the allowed electron energy levels) is characterised by an energy gap between the valence band and conduction band in which there are no permissable

*Mathematical Sciences, University of Southampton, Southampton SO17 1BJ.

electron states. This is illustrated in figure 1. Although it is usual to think of the valence and conduction bands being continuous in energy they are in fact comprised of a large number of states, separated by very small energies. We therefore need to account for the density of states per unit energy in both conduction and valence bands. These are defined as follows

$$\begin{array}{l}\text{Number of states per unit volume in conduction}\\ \text{band with energy between } E \text{ and } E+\delta E\end{array} \approx g_c(E)\delta E, \tag{1}$$

$$\begin{array}{l}\text{Number of states per unit volume in valence}\\ \text{band with energy between } E \text{ and } E+\delta E\end{array} \approx g_v(E)\delta E, \tag{2}$$

where δE is a small increment in energy. Now since electrons are fermions they obey a Fermi-Dirac distribution at equilibrium. Thus the probability of a particular state, with energy E, being occupied at temperature T is given by

$$\mathcal{P}(E,T) = \frac{1}{1+\exp((E-E_f)/k_BT)} \tag{3}$$

where E_f is a constant, termed the Fermi level, which is chosen so that the sum of the probabilities over all possible states gives the total number of electrons in the system. Assuming we know the density of states we can now work out what the electron density in the valence band n_v and in the conduction bands n_c using (1)-(3)

$$n_c = \int_{E_c}^{\infty} \frac{g_c(E)}{1+\exp((E-E_f))/k_BT)} dE, \tag{4}$$

$$n_v = \int_{-\infty}^{E_v} \frac{g_v(E)}{1+\exp((E-E_f)/k_BT)} dE. \tag{5}$$

Note that the conduction band levels all lie above E_c while the valence band levels all lie below E_v.

The Boltzmann approximation. In a typical semiconductor, at room temperature, the valence band is very nearly full (*i.e.* almost all the states are occupied) while the conduction band levels are almost all empty (*i.e.* almost all the states are unoccupied). This is a consequence of the energy gap $E_g = E_c - E_v$ being large with respect to k_BT (*i.e.* $E_g \gg k_BT$) and there being almost the right number of electrons to exactly fill the valence states. The latter is equivalent to the following conditions on the Fermi level

$$\frac{E_c - E_f}{k_BT} \gg 1 \quad \text{and} \quad \frac{E_f - E_v}{k_BT} \gg 1. \tag{6}$$

It follows that the integrand in (5) is almost $g_v(E)$ and this suggests that, rather than counting the number of electrons in the valence band, we should count the number of empty

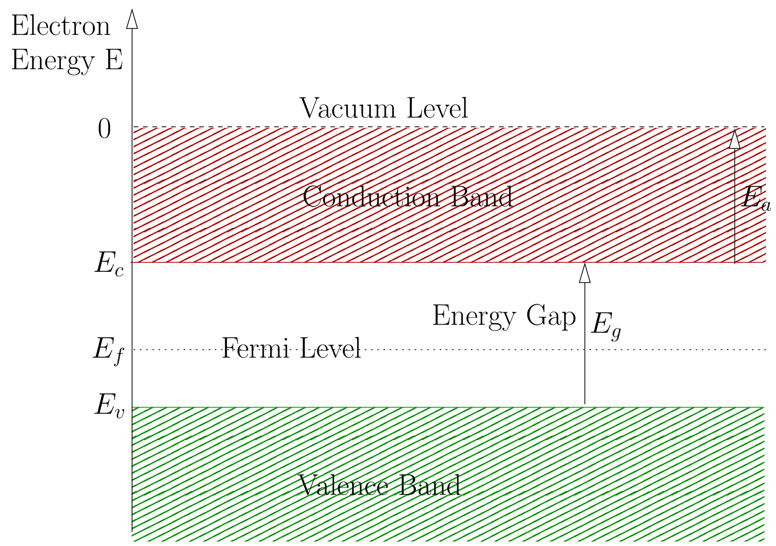

Figure 1: The allowed electron energy levels in a semiconductor (at zero electric potential). Here E_{ea} is the ionisation potential and E_g the gap energy.

states. These empty states are, by convention, called holes and their density p is given by

$$\begin{aligned} p &= \int_{-\infty}^{E_v} g_v(E)dE - \int_{-\infty}^{E_v} \frac{g_v(E)}{1+\exp((E-E_f)/k_BT)}dE \\ &= \int_{-\infty}^{E_v} \frac{g_v(E)\exp((E-E_f)/k_BT)}{1+\exp((E-E_f)/k_BT)}dE \\ &= \int_{-\infty}^{E_v} \frac{g_v(E)}{1+\exp((E_f-E)/k_BT)}dE. \end{aligned}$$

Since the exponential $\exp((E_f - E)/k_BT)$ is very large where the condition (6) is satisfied we can accurately approximate the expression for p by

$$p \sim \int_{-\infty}^{E_v} g_v(E)\exp(-(E_f-E)/k_BT)dE.$$

On making the subsitition $E = E_v - \mathcal{V}$ we obtain the expression

$$p \sim \exp\left(-\frac{E_f - E_v}{k_BT}\right)\int_0^\infty g_v(E_v-\mathcal{V})\exp\left(-\frac{\mathcal{V}}{k_BT}\right)d\mathcal{V}. \tag{7}$$

In a similar fashion we may approximate n_c (the number density of electrons in the conduction band), since from (6) $\exp((E-E_f)/k_BT)$ is very large. On dropping the subscript from n_c we find

$$n \sim \int_{E_c}^{\infty} g_c(E)\exp(-(E-E_f))/k_BT)dE,$$

Then, on making the subsitition $E = E_c + \mathcal{U}$, we obtain the expression

$$n \sim \exp\left(-\frac{E_c - E_f}{k_B T}\right) \int_0^\infty g_c(E_c + \mathcal{U}) \exp\left(-\frac{\mathcal{U}}{k_B T}\right) d\mathcal{U}. \tag{8}$$

It is more usual to express these approximations for p and n in the form

$$p = N_v(T) \exp\left(-\frac{E_f - E_v}{k_B T}\right) \quad \text{and} \quad n = N_c(T) \exp\left(-\frac{E_c - E_f}{k_B T}\right) \tag{9}$$

where

$$N_c(T) = \int_0^\infty g_c(E_c + \mathcal{U}) \exp\left(-\frac{\mathcal{U}}{k_B T}\right) d\mathcal{U} \tag{10}$$

$$N_v(T) = \int_0^\infty g_v(E_v - \mathcal{V}) \exp\left(-\frac{\mathcal{V}}{k_B T}\right) d\mathcal{V}, \tag{11}$$

where N_c and N_v are termed the effective conduction band and valence band densities of states, respectively. If, as is usual, the density of states in the vicinity of the conduction band and valence band are approximately parabolic, *i.e.* $g_c(E_c + \mathcal{U}) \sim \Upsilon_c \mathcal{U}^2$ for $\mathcal{U} > 0$ and $g_v(E_v - \mathcal{V}) \sim \Upsilon_v \mathcal{V}^2$ for $\mathcal{V} > 0$ for some constants Υ_c and Υ_v, we can evaluate these integrals explicitly.

The intrinsic carrier density. A consequence of the Boltzmann approximation is that product np is independent of the Fermi level, that is

$$np = N_c(T) N_v(T) \exp\left(-\frac{E_g}{k_B T}\right) \quad \text{where} \quad E_g = E_c - E_v,$$

and E_g is termed the energy gap. It is perhaps more usual to write this relation in terms of the intrinsic carrier density n_i defined by

$$np = n_i^2 \quad \text{where} \quad n_i = (N_c(T) N_v(T))^{1/2} \exp\left(-\frac{E_g}{2k_B T}\right). \tag{12}$$

The reason for doing this is that in an intrinsic (un-doped) semiconductor we have (at equilibrium)

$$n = p = n_i.$$

Doping. Impurities in the semiconducting crystal can lead to dramatic changes in its electrical properties. Foreign atoms in its crystal lattice may preferentially donate electrons to the conduction band (n-type doping) or grab them from the valence band (p-type doping) leading to an imbalance between the number of holes and electrons. In the case of n-type doping this leads to a stationary net positive charge (the impurity atom loses one of its electrons) while in the case of p-type doping this leads to a stationary net negative charge (the impurity atom gains an electron). If these impurities occur in sufficient quantities they may significantly increase the number of charge carriers above the intrinsic carrier density. This is a consequence of the net charge in the material having to remain close to zero (charge neutrality). Thus if N_d denotes the density of donor impurities (fixed positive charges) while N_a denotes the density of acceptor impurities (fixed negative charges) the condition of charge neutrality implies

$$q(p - n + N_d - N_a) = 0. \tag{13}$$

Hence if there are very few acceptor impurities and a large number of donor impurities with $N_d \gg n_i$ we expect (from (12)-(13)) that

$$n \approx N_d \qquad p \approx \frac{n_i^2}{N_d}$$

Such a material is termed n-doped because it has very many more conduction electrons than holes. Alternatively if there are very few donor impurities and a large number of acceptor impurities with $N_a \gg n_i$ we expect (from (12)-(13)) that

$$p \approx N_a \qquad n \approx \frac{n_i^2}{N_a}$$

Such a material is termed p-doped because it has very many more holes than conduction electrons.

Doping can be used to enhance the conductivity of a semiconductor because a doped materical typically has many more charge carriers than an undoped (or intrinsic) material. It can also be used as a mechanism to separate holes from electrons by p-doping a semiconductor in one region and n-doping it in another. This sets up a potential difference between the two regions that acts to drive holes into the p-doped region and electrons into the n-doped region and can thus be used to separate solar generated charge carrier pairs in a photovoltaic device.

3 Probabilistic and Drift-Diffusion models of charge transport

We begin by considering models for a particle hopping on a lattice of energy wells and show that in a certain limit the probability of finding the particle in a given well at a particuar time can be approximated by a diffusion equation. We generalise to a particle hopping on a lattice with an externally applied potential and show that its probability density satisfies an drift diffusion equation. This analysis then applied to the motion of charge carriers in a semiconductor. We finish by applying the drift diffusion model of charge carrier transport in a semiconductor to a photovoltaic device formed from a p-n junction and show that its current voltage curve is identical to that given by a simple equivalent circuit model of a solar cell.

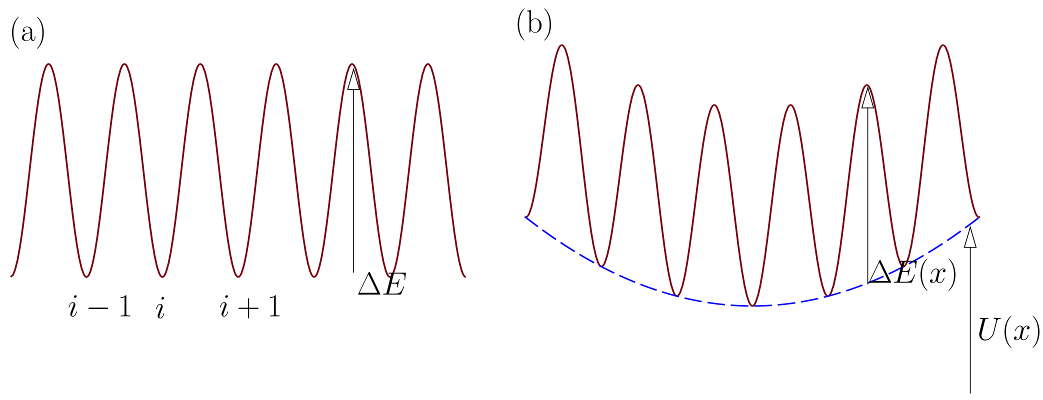

Figure 2: The energy landscape of the particle for (a) a lattice with no applied potential and (b) a lattice with applied potential $U(x)$.

3.1 Rate equations for a particle hopping on a lattice

We start by considering a particle in one dimensional periodic energy landscape as depicted in figure 2(a). This is characterised by energy wells separated by intervening energy peaks. The particle sitting in one of these wells will, from time to time, have sufficient thermal energy to traverse the energy maximum dividing this well from one of its neighbours. The rate at which this "hopping" process occurs is given by the Arrhenius equation

$$\begin{aligned}
\mathcal{P}_{i\to i+1} = \text{Probability particle moves from well } i \text{ to } i+1 \text{ per unit time} &= K \exp\left(-\frac{\Delta E}{k_B T}\right) \\
\mathcal{P}_{i\to i-1} = \text{Probability particle moves from well } i \text{ to } i-1 \text{ per unit time} &= K \exp\left(-\frac{\Delta E}{k_B T}\right)
\end{aligned}$$

where k_B is Boltzmann's constant; T is absolute temperature; ΔE is the energy difference between two neighbouring wells; and K a phenomenological rate constant.

The Dynamic Monte-Carlo method. We could at this stage simulate the motion of the particle simply by getting a computer to throw dice in order to decide what it does. The simplest way to do this is to pick a small time step Δt with the property that $\Delta t K \exp\left(-\frac{\Delta E}{k_B T}\right) \ll 1$ (this means that in any given time step the chance of the particle moving is small). Then to pick a random number R evenly distributed between 0 and 1 and implement the procedure

$$\begin{aligned}
&\text{If} \quad 0 \le R < \Delta t K \exp\left(-\frac{\Delta E}{k_B T}\right) && \text{then} \quad i(t+\Delta t) = i(t) - 1, \\
&\text{If} \quad \Delta t K \exp\left(-\frac{\Delta E}{k_B T}\right) \le R < 2\Delta t K \exp\left(-\frac{\Delta E}{k_B T}\right) && \text{then} \quad i(t+\Delta t) = i(t) + 1, \\
&\text{If} \quad 2\Delta t K \exp\left(-\frac{\Delta E}{k_B T}\right) \le R \le 1 && \text{then} \quad i(t+\Delta t) = i(t).
\end{aligned}$$

This is an extremely easy procedure to implement but it is numerically inefficient because for most of the time steps nothing happens and the particle remains where it is (to put it another way it takes very little of your time to program but you may have to wait a long time for the results if you are investigating a more complex problem than this).

The Gillespie Algorithm. It is, as remarked on above, computationally wasteful to take very small timesteps in which nothing happens. An alternative, more efficient procedure, is to calculate the randomly distributed time T between the last event and the next one and then to calculate what the next event is based on the relative probabilities of the possible events. Thus in the case of our previous example of a particle hopping on a one-dimensional lattice the next event is either that the particle hops right (with probablity per unit time $\mathcal{P}_{i\to i+1}$) or that it hops left (with probablity per unit time $\mathcal{P}_{i\to i-1}$). So the

$$\left(\begin{array}{l}\text{Probability of an event}\\ \text{occuring in the}\\ \text{infinitessimal interval}\\ T<t<T+\delta t\end{array}\right) = \mathcal{P}_{tot}\delta t \qquad \text{where} \quad \mathcal{P}_{tot} = \mathcal{P}_{i\to i-1} + \mathcal{P}_{i\to i+1}. \tag{14}$$

Given that the last event occurred at $t = 0$ the probability of the next event occurring in the infinitessimal time interval $T < t < T + \delta T$ is

$$\left(\begin{array}{l}\text{Probability next event}\\ \text{occurs in}\\ T<t<T+\delta t\end{array}\right) = \left(\begin{array}{l}\text{Probability no events}\\ \text{occurs in interval}\\ 0<t<T\end{array}\right) \times \left(\begin{array}{l}\text{Probability an event}\\ \text{occurs in interval}\\ T<t<T+\delta t\end{array}\right), \tag{15}$$

In order to evaluate this expression we need to calculate the probability that no events occur in $0 < t < T$ (we have already calculated the probability that an even occurs in $T < t < T+\delta t$ in (14)). In order to do this we subdivide the interval $0 < t < T$ into a large number, N, of very small intervals of size Δt such that $N\Delta t = T$. An approximate expression for this quantity is then given by

$$\begin{aligned}\left(\begin{array}{l}\text{Prob. no events}\\ \text{in interval}\\ 0<t<T\end{array}\right) &\approx \left(\begin{array}{l}\text{Prob. no events}\\ \text{in interval}\\ 0<t<\Delta t\end{array}\right) \times \left(\begin{array}{l}\text{Prob. no events}\\ \text{in interval}\\ \Delta t<t<2\Delta t\end{array}\right) \times \cdots \\ &\qquad \cdots \times \left(\begin{array}{l}\text{Prob. no events}\\ \text{in interval}\\ 0<t<N\Delta t\end{array}\right) \\ &= (1-\mathcal{P}_{tot}\Delta t)\times(1-\mathcal{P}_{tot}\Delta t)\times\cdots\times(1-\mathcal{P}_{tot}\Delta t), \\ &= (1-\mathcal{P}_{tot}\Delta t)^N.\end{aligned}$$

Taking the (natural) logarithm of this expression we find that

$$\begin{aligned}
\log_e \begin{pmatrix} \text{Prob. no events} \\ \text{in interval} \\ 0 < t < N\Delta t \end{pmatrix} &\approx N \log_e(1 - \mathcal{P}_{tot}\Delta t) \\
&\approx -N\mathcal{P}_{tot}\Delta t \quad \text{since} \quad \mathcal{P}_{tot}\Delta t \ll 1 \\
&= -\mathcal{P}_{tot}T.
\end{aligned}$$

It follows that

$$\begin{pmatrix} \text{Prob. no events} \\ \text{in interval} \\ 0 < t < N\Delta t \end{pmatrix} = \exp(-\mathcal{P}_{tot}T).$$

Substituting this result, together with (14), back into (15) we obtain the desired result

$$\begin{pmatrix} \text{Probability next event} \\ \text{occurs in} \\ T < t < T + \delta t \end{pmatrix} = \mathcal{P}_{tot}\delta t \exp(-\mathcal{P}_{tot}T). \tag{16}$$

How then should we use the result (16) to randomly select the time T at which the next even occurs? Recall that it is easy to get a computer to pick a (uniformly distributed) random number R with value between 0 and 1. Furthermore the probability that the next event occurs in the interval $0 < t < T$ is a random variable that is distributed between 0 and 1 (*i.e.* it is 0 when $T = 0$ and 1 when $T = \infty$). Indeed by integrating the expression (16) we find that

$$\begin{aligned}
\begin{pmatrix} \text{Probability next event} \\ \text{occurs in range} \\ 0 < t < T \end{pmatrix} &= \int_0^T \mathcal{P}_{tot} \exp(-\mathcal{P}_{tot}t)dt & (17) \\
&= [-\exp(-\mathcal{P}_{tot}t)]_0^T & (18) \\
&= 1 - \exp(-\mathcal{P}_{tot}T). & (19)
\end{aligned}$$

In order to calculate the time T to the next event we therefore identify $1 - \exp(-\mathcal{P}_{tot}T)$ with the uniformly distributed variable R, with range $0 < R < 1$, by writing

$$1 - \exp(-\mathcal{P}_{tot}T) = R$$

from which we calculate T as

$$T = -\frac{1}{\mathcal{P}_{tot}} \log_e(1 - R). \tag{20}$$

Once we have calculated the time T to the next event (using the uniformly distributed variable R with range $0 < R < 1$) we need to work out which of the two possible events (the particle moves left or the particle moves right) occurs at this time. This is easy to do since the relative probabilities of these two events are

$$\begin{pmatrix} \text{Relative probability} \\ \text{particle moves right} \\ i \to i+1 \end{pmatrix} = \frac{\mathcal{P}_{i\to i+1}}{\mathcal{P}_{tot}}, \qquad \begin{pmatrix} \text{Relative probability} \\ \text{particle moves left} \\ i \to i-1 \end{pmatrix} = \frac{\mathcal{P}_{i\to i-1}}{\mathcal{P}_{tot}}.$$

and the relative probabilites clearly add to 1. We therefore calculate which of the two possibilities occurs (in the next event) by getting the computer to pick a second (uniformly distributed) random number r, with value between 0 and 1, and

$$\begin{array}{llll} \text{moving the particle moves to the right} & (i.e.\ i \to i+1) & \text{if} & 0 < r < \dfrac{\mathcal{P}_{i\to i+1}}{\mathcal{P}_{tot}}, \\ \text{moving the particle moves to the left} & (i.e.\ i \to i-1) & \text{if} & \dfrac{\mathcal{P}_{i\to i+1}}{\mathcal{P}_{tot}} < r < 1. \end{array} \tag{21}$$

The Algorithm. To recap the Gillespie Algorithm for a particle hopping on a one-dimensional lattice consists of the steps

1. Use the computer to pick a (uniformly distributed) random number R with value between 0 and 1.

2. Calculate the time T to the next event from the value of this random number with the formula

$$T = -\frac{1}{\mathcal{P}_{tot}} \log_e(1 - R).$$

3. Use the computer to pick a second (uniformly distributed) random number r with value between 0 and 1.

4. Determine which of the two events occur from the value of r based on the criterion

$$\begin{array}{llll} \text{particle moves to the right} & (i.e.\ i \to i+1) & \text{if} & 0 < r < \dfrac{\mathcal{P}_{i\to i+1}}{\mathcal{P}_{tot}}, \\ \text{particle moves to the left} & (i.e.\ i \to i-1) & \text{if} & \dfrac{\mathcal{P}_{i\to i+1}}{\mathcal{P}_{tot}} < r < 1. \end{array} \tag{22}$$

5. Repeat.

Generalisation to more complicated systems. With a bit of common sense this algortihm can be generalised to more complicated random systems where there are more than two possible events. For example the motion of a particle on a square lattice; here there are 4 possible events (i) particle moves left, (ii) particle moves right, (iii) particle moves up, (iv) particle moves down. Or for example N particles moving on a square lattice; here there are $4N$ possible events corresponding to each particle moving in 1 of 4 directions. Or if we are interested in carrying out a Monte-Carlo simulation of a solar cell we need to account for the motion of a large number of holes and electrons on a lattice together with generation events (creation of a hole and electron) and recombination events (destruction of a hole and electron).

3.2 A probabilistic approach to particle hopping.

Rather than trying to directly simulate the motion of the particle it is often sufficient to track the probability of the particle being at any particular lattice point as a function of time. With this in mind we define

$$P_i(t) = \begin{array}{l}\text{Probability particle is}\\ \text{in well } i \text{ at time } t\end{array}$$

and attempt to write down a series of coupled ordinary differential equations (ODEs) in time for the probabilities $P_i(t)$ by considering the evolution of the system over an infinitesimal time interval δt. Suppose that we know what the probability of finding the particle is at any of the lattice points at time t (that is we know all the $P_i(t)$ for all values i) then the probability that the particle is in the i'th well at time $t + \delta t$ is given approximately by

$$\begin{aligned} P_i(t+\delta t) \;\approx\; & P_i(t) \times \left(\begin{array}{l}\text{Probability particle remains}\\ \text{in well } i \text{ over interval } \delta t\end{array} \right) \\ & + P_{i+1}(t) \times \left(\begin{array}{l}\text{Probability particle jumps from}\\ \text{well } i+1 \text{ to well } i \text{ in interval } \delta t\end{array} \right) \\ & + P_{i-1}(t) \times \left(\begin{array}{l}\text{Probability particle jumps from}\\ \text{well } i-1 \text{ to well } i \text{ in interval } \delta t\end{array} \right). \end{aligned} \tag{23}$$

It follows that

$$\begin{aligned} P_i(t+\delta t) \;&\approx\; P_i(t)\left(1 - 2\delta t K \exp\left(-\frac{\Delta E}{k_B T}\right)\right) \\ &\quad + P_{i+1}(t)\delta t K \exp\left(-\frac{\Delta E}{k_B T}\right) + P_{i-1}(t)\delta t K \exp\left(-\frac{\Delta E}{k_B T}\right) \\ \Longrightarrow \frac{P_i(t+\delta t) - P_i(t)}{\delta t} \;&=\; K \exp\left(-\frac{\Delta E}{k_B T}\right)\left(P_{i+1}(t) - 2P_i(t) + P_{i-1}(t)\right). \end{aligned}$$

Taking the limit $\delta t \to 0$

$$\frac{dP_i}{dt} = K \exp\left(-\frac{\Delta E}{k_B T}\right)\left(P_{i+1}(t) - 2P_i(t) + P_{i-1}(t)\right). \tag{24}$$

3.3 Derivation of a diffusion equation for particle hopping.

In practise we might be thinking of applying this approach to a problem in which there are a very large number of lattice sites: for example to an electron hopping along a line of atoms in a crystal. In such scenarios it is not sensible to track of the probability on all lattice sites. Furthermore, since the probability at neighbouring lattice sites is likely to be very similar, we can seek to replace the probabilities $P_i(t)$ on the lattice sites by a continuum representation of the probability $p(x,t)$, where we suppose x to be a spatial representation of the site's position. Thus if the lattice sites are evenly spaced a distance δx apart we can write $x = i\delta x$. It is tempting to identify $p(x,t)$ directly with $P_i(t)$ by writing $p(i\delta x, t) = P_i(t)$ but rather than do this we instead choose to define $p(x,t)$ as a probability density by writing

$$p(i\delta x, t) = \frac{P_i(t)}{\delta x}$$

[Aside: You might be worried about the dimensions here because probability densities usually have dimensions of L^{-3} and $p(x,t)$ seems to have dimensions of L^{-1}. However we note that if we were considering a three-dimensional problem we would have defined $p(x,y,z,t)$ by

$$p(i\delta x, j\delta x, k\delta x, t) = \frac{P_{ijk}(t)}{\delta x^3} \tag{25}$$

which has the right dimensions and so if we are thinking of this as being a one-dimensional representation of a 3-d problem the units of $P_i(t)$ really should be L^{-2} for the obvious reasons.]

Having defined $p(x,t)$ by (25) and substituted this into (24) we retrieve

$$\frac{\partial p}{\partial t}(x,t) = \left[K\delta x^2 \exp\left(-\frac{\Delta E}{k_B T}\right)\right] \frac{p(x+\delta x,t) - 2p(x,t) + p(x-\delta x,t)}{\delta x^2}$$

If we now Taylor expand $p(x+\delta x,t)$ and $p(x-\delta x,t)$ about x (assuming that δx is sufficiently small to allow us to do so) we find that, to a leading order approximation, that $p(x,t)$ satisfies a diffusion equation

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2}, \tag{26}$$

where

$$D = \left[K\delta x^2 \exp\left(-\frac{\Delta E}{k_B T}\right)\right]. \tag{27}$$

Interpretation in terms of particle number density (concentration). Thus far we have assumed only a single particle hopping on a lattice but in reality we are much more often concerned with a large number of particles on a lattice (*e.g.* free electrons in an organic semiconductor or excitons). This does not present a problem provided that the particles do not interact with each other via, for example, Coulomb forces or exclusion effects. We will leave the treatment of interaction via Coulomb forces until §3.4 and note here that exclusion effects only become significant when the number of particles is comparable to the number of lattice sites so that there is competition between particles for available lattice sites. Hence, provided the particles do not interact via a force and their number is much smaller than that of the lattice sites, we can treat them as effectively independent. This means each of the N particles has a probability density $p^{(n)}(x,t)$ (for $n = 1, 2, \cdots, N$) which is, to a good approximation, independent of that for all the other particles. It follows that the particle number density $c(x,t)$ defined by

$$c(x,t) = \sum_{n=1}^{N} p^{(n)}(x,t),$$

obeys the same equation that the probability distribution for each particle obeys (*i.e.* (26)). Thus

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}, \tag{28}$$

3.4 Description of a particle hopping on a lattice in an applied potential

Semiconductors contain relatively large numbers of mobile charged particles (free electrons and holes) which, since they are charged, interact with each other via the Coulomb potential. In order to treat this type of scenario we consider a single particle moving in a potential comprised of a quasi-periodic part (as previously) and a smooth part $U(x)$ (the applied potential); this is illustrated in figure 2(b). In order to make the analysis more general we allow the energy barrier between the wells (in the absence of the applied potential) ΔE to vary smoothly with x; the implicit assumption being that ΔE varies by a small amount between neighbouring peaks but can vary significantly over many peaks. We denote the energy barrier on passing from well i to well $i+1$ by $\Delta E_{i\to i+1}$ and that on passing from well i to well $i-1$ by $\Delta E_{i\to i-1}$. In the presence of the applied potential $U(x)$ the appropriate expressions for these quantities are

$$\Delta E_{i\to i+1} = U\left(x+\frac{\delta x}{2}\right) - U(x) + \Delta E\left(x+\frac{\delta x}{2}\right), \tag{29}$$

$$\Delta E_{i\to i-1} = U\left(x-\frac{\delta x}{2}\right) - U(x) + \Delta E\left(x-\frac{\delta x}{2}\right), \tag{30}$$

where, as before $x = i\delta x$. In a similar fashion

$$\Delta E_{i+1\to i} = U\left(x+\frac{\delta x}{2}\right) - U(x+\delta x) + \Delta E\left(x+\frac{\delta x}{2}\right), \tag{31}$$

$$\Delta E_{i-1\to i} = U\left(x-\frac{\delta x}{2}\right) - U(x-\delta x) + \Delta E\left(x-\frac{\delta x}{2}\right). \tag{32}$$

Once again we consider a small time interval δt and formulate a set of coupled difference equations for $P_i(t+\delta t)$ (the probability of the particle being in well i at time $t+\delta t$) as described in equation (23). The rate at which the particle jumps from one well to a neighbouring well are (as before) determined by the energy barrier that it has to surmount, through Arrhenius' Law. We can thus translate (23) into the following difference equation

$$\begin{aligned} P_i(t+\delta t) \approx{}& P_i(t)\left(1-\delta t K\left(\exp\left(-\frac{\Delta E_{i\to i+1}}{k_B T}\right)+\exp\left(-\frac{\Delta E_{i\to i-1}}{k_B T}\right)\right)\right) \\ &+P_{i+1}(t)\delta t K\exp\left(-\frac{\Delta E_{i+1\to i}}{k_B T}\right)+P_{i-1}(t)\delta t K\exp\left(-\frac{\Delta E_{i-1\to i}}{k_B T}\right). \end{aligned}$$

On rearranging this equation and taking the limit as $\delta t \to 0$ we obtain a set of coupled ODEs (exactly as before), namely

$$\begin{aligned} \frac{dP_i}{dt} ={}& K\left[P_{i+1}(t)\exp\left(-\frac{\Delta E_{i+1\to i}}{k_B T}\right)+P_{i-1}(t)\exp\left(-\frac{\Delta E_{i-1\to i}}{k_B T}\right)\right. \\ &\left.-P_i(t)\left(\exp\left(-\frac{\Delta E_{i\to i+1}}{k_B T}\right)+\exp\left(-\frac{\Delta E_{i\to i-1}}{k_B T}\right)\right)\right]. \end{aligned} \tag{33}$$

In order to obtain a partial differential equation for the probability density $p(x,t)$ (as defined in (25)) we proceed as before and expand in powers of δx (having first substituted for the energy barriers from (29)-(32)). As this is a rather lengthy procedure we won't give the details here but we note that that the calculation is of a similar nature to that used to derive (26) and that the resulting PDE is

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial x}\left(D(x)\left(\frac{\partial p}{\partial x}+\frac{p}{k_B T}\frac{\partial U}{\partial x}\right)\right), \tag{34}$$

where

$$D(x) = \left[K\delta x^2 \exp\left(-\frac{\Delta E(x)}{k_B T}\right)\right].$$

Using arguments analogous to those presented at the end of §3.3 we can show that the same equation is satisfied by the particle number density $c(x,t)$, *i.e.*

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\left(D(x)\left(\frac{\partial c}{\partial x}+\frac{c}{k_B T}\frac{\partial U}{\partial x}\right)\right), \tag{35}$$

in a system in which the number of particles is much smaller than the number of lattice sites.

Generalisation to higher dimensions. It is straightforward (but slightly messy) to show that the one dimensional drift diffusion equation generalises to three-dimensions to give, as might be expected,

$$\frac{\partial c}{\partial t} = \nabla \cdot \left(D(\boldsymbol{x}) \left(\nabla c + \frac{c}{k_B T} \nabla U \right) \right), \tag{36}$$

3.5 Other processes giving rise to advection-diffusion models

The above description of particle motion in a periodic energy landscape could easily be adapted to free electrons in the LUMO (lowest unoccupied molecular orbital) of an organic semiconductor (or equally to a hole in the HOMO of an organic semiconductor). However it does not apply to conduction electrons in an inorganic semiconductor which occupy a genuine conduction band and are therefore free to move largely unhindered through the crystal lattice. Nevertheless the same drift-diffusion equation, derived above in (34), applies to the motion of electrons in a conduction band and holes in a valence band except that here the diffusive character of the motion can be attributed to electron scattering in the crystal lattice. The derivation of the drift-diffusion equations in an inorganic semiconductor is outlined in Nelson (chapter 3) [7].[1] This derivation is considerably more involved than the one given above for hopping processes. It starts (I) from a quantum mechanical description of the wavefunction of an electron in the conduction band of a semiconducting lattice (using Bloch wave functions), (II) assumes that electrons are in quasi thermal equilibrium at all points in space (*i.e.* scattering of electrons within the crystal is sufficiently fast that the electrons are distributed [across energy levels of the conduction band] according to a Boltzmann distribution) and (III), having made this assumption, uses the Boltzmann Transport equation to calculate the electron current in terms of the electric field and the spatial distribution of electrons.

4 Drift diffuion models of charge transport in semiconductors

Charge transport in a semiconductor is characterised by the motion of free electrons (in the conduction band [inorganic] or LUMO [organic]) and holes (in the valence band [inorganic] or HOMO [organic]). Both electrons and holes are charged and therefore interact via the Coulomb force. In order to keep track of these interactions it is usual to define an 'averaged' electric potential ϕ which satisfies Poisson's equation

$$\nabla \cdot (\varepsilon \nabla \phi) = -\rho(\boldsymbol{x}, t) \tag{37}$$

where ε is the permittivity. The charge density ρ is typically approximated in terms of p, n and C_{fix} which are the hole, free electron and the fixed charge densities (or in other

[1] See also Feynman Volume III [1] Chapters 13-14.

words $C_{fix} = N_d - N_a$ is the net doping density), respectively. The relation given by this approximation is

$$\rho(\boldsymbol{x}, t) = q(p(\boldsymbol{x}, t) - n(\boldsymbol{x}, t) + C_{fix}) \tag{38}$$

where q is the charge on a proton. In making this approximation we have smeared out the effects of the individual point charges by replacing the delta-functions in the charge density with a smooth function $q(p(\boldsymbol{x}, t) - n(\boldsymbol{x}, t) + C_{fix})$. Where opposite charges do not tend to pair up this is a sensible approach to adopt. Indeed it works well for semiconductors however, there are systems, such as electrolytes, where this approximation is not particularly good and alternative approaches have to be adopted.

Assuming that there is very little charge recombination or generation the resulting equations for n, p and ϕ can be found by combining (37) with (38), writing $c = p$ and $U = q\phi$ in (36) and writing $c = n$ and $U = -q\phi$ in (36) to obtain the set of coupled PDEs

$$\frac{\partial p}{\partial t} = \nabla \cdot \left(D_p \left(\nabla p + \frac{q}{k_B T} p \nabla \phi \right) \right), \tag{39}$$

$$\frac{\partial n}{\partial t} = \nabla \cdot \left(D_n \left(\nabla n - \frac{q}{k_B T} n \nabla \phi \right) \right), \tag{40}$$

$$\nabla \cdot (\varepsilon \nabla \phi) = q(n - p - C_{fix}). \tag{41}$$

Here D_p and D_n are the diffusivity of holes and mobile electrons, respectively.

Remarks. It is notable that there are 3 PDEs for the 3 variables p, n and ϕ and we therefore seem to have the right number of equations for unknowns. There are however a number of things missing from the physics. Firstly we have entirely neglected charge generation and recombination which are key ingredients if we want to model a solar cell. Secondly we have implicitly assumed that we are considering only a single materialn when in fact we may want to investigate devices with junctions between a number of semiconducting materials.

4.1 Currents, fluxes and carrier concentration

In order to aid the physical interpretation of their solutions it is helpful to recast the drift-diffusion equations for free electrons and holes in terms of alternative variables (*e.g.* quasi-Fermi levels). We begin re-expressing (36) (or alternatively (39)-(40)) in terms of conservation laws for the carrier concentrations. On writing $c = p$ in (36) and subsequently writing $c = n$ in (36) and re-arranging the resulting equations we find

$$\frac{\partial p}{\partial t} + \nabla \cdot (p \boldsymbol{v}_p) = 0 \quad \text{where} \quad \boldsymbol{v}_p = -\frac{D_p}{k_B T} \nabla \left(k_B T \ln \left(\frac{p}{c_p^*} \right) + U_p \right), \tag{42}$$

$$\frac{\partial n}{\partial t} + \nabla \cdot (n \boldsymbol{v}_n) = 0 \quad \text{where} \quad \boldsymbol{v}_n = -\frac{D_n}{k_B T} \nabla \left(k_B T \ln \left(\frac{n}{c_n^*} \right) + U_n \right). \tag{43}$$

Here it is not immediately clear what the reference concentrations c_p^* and c_n^* should be, and it does not matter provided we only consider a single homogenous material in which these concentrations are constant so that their gradients are zero. In fact it turns out (for sound statistical mechanical reasons) that we should identify c_p^* and c_n^* with the effective densities of states in the valence and conduction bands, respectively (see (11)) so that

$$c_p^* = N_v \quad \text{and} \quad c_n^* = N_c$$

We may identify (42a) and (43a) as conservation equations for p and n respectively in which $\boldsymbol{v}_p$ and $\boldsymbol{v}_n$ represent the average velocities of holes and electrons respectively. Equations (42b) and (43b) for the average velocities can then be re-expressed in terms of thermodynamic quantities as

$$\boldsymbol{v}_p = -\frac{D_p}{k_B T}\nabla\mu_p \quad \text{where} \quad \mu_p = k_B T \ln\left(\frac{p}{N_v}\right) + U_p, \tag{44}$$

$$\boldsymbol{v}_n = -\frac{D_n}{k_B T}\nabla\mu_n \quad \text{where} \quad \mu_n = k_B T \ln\left(\frac{n}{N_c}\right) + U_n. \tag{45}$$

Here μ_p and μ_n are the electrochemical potentials of holes and free electrons, respectively.

The hole current density $\boldsymbol{j}_p$ (*i.e.* that portion of the current carried by holes) and the electron current density $\boldsymbol{j}_n$ are related to the average carrrier velocities by $\boldsymbol{j}_p = qp\boldsymbol{v}_p$ and $\boldsymbol{j}_n = -qp\boldsymbol{v}_n$ from which it follows that

$$\boldsymbol{j}_p = -\frac{qD_p}{k_B T} p\nabla\mu_p \quad \text{and} \quad \boldsymbol{j}_n = \frac{qD_n}{k_B T} n\nabla\mu_n.$$

In the solid state physics literature these quantities are more usually expressed in terms of carrier mobilities and quasi Fermi levels in the form

$$\boldsymbol{j}_p = M_p p\nabla E_{F_p} \quad \text{and} \quad \boldsymbol{j}_n = M_n n\nabla E_{F_n}. \tag{46}$$

Here the electron and hole mobilities (M_n and M_p, respectively) and the electron and hole quasi Fermi levels (E_{F_n} and E_{F_p}, respectively) are defined by

$$M_n = \frac{qD_n}{k_B T}, \quad M_p = \frac{qD_p}{k_B T}, \quad E_{F_n} = \mu_n, \quad E_{F_p} = -\mu_p. \tag{47}$$

Electron and hole potentials and quasi-Fermi levels. We now need to specify the electron and hole potentials (U_n and U_p respectively) more precisely. We begin by considering the mobile electrons. These electrons lie close to the lower edge of the conduction band and so in the absence of an electric field (*i.e.* zero electric potential) their energy is $-E_{ea}$, where E_{ea} is the electron affinity; *i.e.* the energy required to remove an electron from the conduction band edge to the vacuum level (see figure 1). The electrostatic part of the potential is $-q\phi$. Adding these two contributions together gives

$$U_n(x) = -E_{ea}(x) - q\phi(x). \tag{48}$$

The holes lie close to the upper edge of the valence band and in order to calculate their potential we calculate the potential of an electron lying there and note that the potential of a hole (*i.e.* a lack of an electron) is just minus that of the electron it replaces. It follows that

$$U_p(x) = (E_{ea}(x) + E_g(x)) + q\phi(x), \tag{49}$$

where E_g is the energy gap between the valence and conduction band[2] (see figure 1). It is also common to see the electron and hole potentials written in terms of the electron energies (including the effect of the electric potential) at the band edges

$$U_n(x) = E_c(x) \quad \text{and} \quad U_p(x) = -E_v(x). \tag{50}$$

Here E_c and E_v are the energies at the conduction and valence band edges, respectively. It now becomes apparent why E_{F_n} and E_{F_p} are termed quasi-Fermi levels because if we take our definitions of the quasi-Fermi levels (47c-d) and substitute for the chemical potentials μ_p and μ_n from (44b) and (45b) and rewrite the resulting expression in terms of p and n (using (50)) we obtain

$$p = N_v \exp\left(-\frac{E_{F_p} - E_v}{k_B T}\right) \quad \text{and} \quad n = N_c \exp\left(-\frac{E_c - E_{F_n}}{k_B T}\right). \tag{51}$$

These look almost identical to (9) the formulae for the electron and hole denisities at equilibrium except that the Fermi level in the equation of for n has been replaced by the quasi-Fermi level E_{F_n} and the Fermi level in the equation of for p has been replaced by the quasi-Fermi level E_{F_p}. However the formulae (51) dont really help us a lot because we cannot use them to determine the carrier concentrations n and p and it is much better to think of the quasi-Fermi levels as being determined by the carrier concentrations rather than the other way round. With this in mind we use the definitions of the quasi-Fermi levels in (47c-d), (45b) and (44b) together with those of the electron and hole potentials to write

$$E_{F_n} = k_B T \ln\left(\frac{n}{N_c}\right) - E_{ea} - q\phi, \quad E_{F_p} = -k_B T \ln\left(\frac{p}{N_v}\right) - E_{ea} - E_g - q\phi. \tag{52}$$

4.2 Carrier generation and recombination

Thinking back to our treatment of a semiconductor at equilibrium it becomes apparent that we have forgotten something. Our equations for the conservation of electrons (42)-(43) allow only for transport of electrons within a band but no interchange of electrons between the two bands (conduction and valence). Yet when we discussed the equilibrium problem we assumed that there was thermal equilibrium between the electrons in the two bands and so we clearly need to allow jumping of electrons between bands, even if this might be a relatively small effect with respect to other processes. If we don't do this we cannot expect the solution of our

[2] Here $-(E_g + E_{ea})$ is the energy of an electron at the upper edge of the valence band at zero electric potential.

model to eventually settle down to an equilibrium. How should we model this? Equations (42)-(43) are conservation equations for holes and free electrons respectively. Now moving an electron from the valence band to the conduction band results in the creation of a hole (in the valence band) and a free electron (in the conduction band). Similarly a free electron can only jump down from the conduction band into the valence band if there is hole for it to fill and so this process therefore results in the annihilation of a free electron and a hole. In other words motion of electrons between bands results in the creation (or annihilation) of equal number of holes and free electrons. It follows that the only sensible way to alter (42)-(43) is by adding the same term to the right-hand side of both these equations so that

$$\frac{\partial p}{\partial t} + \nabla \cdot (p\boldsymbol{v}_p) = G_{therm}, \qquad \frac{\partial n}{\partial t} + \nabla \cdot (n\boldsymbol{v}_n) = G_{therm}. \tag{53}$$

Here it is to be understood that G_{therm} is the volumetric rate of generation of holes and free electrons and can be either positive (if there is positive net rate of electrons jumping from conduction to valence band) or negative (if there is positive net rate of electrons jumping from valence to conduction band). Furthermore we require that, if the semiconductor is isolated and there are no spatial gradients (so that $\boldsymbol{v}_p = \boldsymbol{v}_n = 0$), the solution to (53) settles to the thermodynamic equilibrium given in (12) for long times (*i.e.* $np \to n_i^2$ as $t \to +\infty$). If we think first about the recombination of free electrons (in the conduction band) with holes (in the valence band) we might reasonably expect that, since both species are rare, the rate limiting step is getting the free electron close enough to the hole so that they can annihilate. The rate at which these two species come within a critical reaction radius per unit volume is clearly proportional to the product of their concentrations np. Putting these facts together we end up with a simple (and correct) model for the thermal generation rate

$$G_{therm} = \gamma(n_i^2 - np). \tag{54}$$

for some material dependent rate constant γ.

Solar generation. Central to the use of semiconductors in solar cell is their ability to absorb light and turn the absorbed energy into charge free-electron hole pairs. Clearly such pairs can only be generated directly if the absorbed photon has more energy than the band gap, which is one of the reasons that the choice of material with the right band gap is so important. If the band gap is large not many photons are turned into free-electron hole pairs but the free-electron hole pairs that are generated have a large potential energy (corresponding to a high open circuit voltage V_{oc} and a small short-circuit current J_{sc}). If on the other hand the band gap is small most photons will generate free-electron hole pairs but these will have low potential energy (corresponding to a high J_{sc} and low V_{oc}). In inorganic semiconductors (where the band gap is usually large) the generation of charge pairs typically occurs directly throughout the bulk of the semiconductor. In organic materials, however, charge pair generation typically occurs in a two stage process. Photons are absorbed in the bulk of the semiconductor to produced an exciton (an excited electron) which can be thought of as a coulombically bound electron hole pair. This exciton can diffuse around but has a relatively short lifetime before it releases its energy as heat. Such excitonic materials can only be used in photovoltaics if they form an interface with another material at which it is favourable for excitons to dissociate into a free-electron in one material and a hole in the other. As this is rather complicated we will stick to the direct form of generation (which incidentally also seems to be that relevant for perovskites). At its crudest level we ought be able to treat such carrier generation in a similar fashion to thermal generation by adding a term $G_{sol}(\boldsymbol{x})$ to the right-hand side of (53). Of course the exact details of the size of G_{sol} and its spatial variation are extremely important and as hinted to above depend on band gap, the radiation spectrum and how much radiation has been absorbed in other parts of the cell before reaching the point $\boldsymbol{x}$. However these are not the focus of this course and so we will not go into these here, and typically will assume G_{sol} is a constant (which can only really be justified in a very thin cell which absorbs only a small fraction of the incident light). It is worth noting that the rate of solar generation (per sufficiently energetic phtoton) is intimately linked to the rate of bimolecular recombination γ. This is because solar generation and bimolecular recombination are essentially the same processes but in reverse. For more information on how these rates should be calculated see Chapter 4 [7].[3]

Recombination via trapped states. Thus far we have assumed that charge pairs only recombine via bimolecular recombination, in which an electron from the conduction band drops direcly into a hole in the valence band; this is represented by the second (negative) term on the right-hand side of (54) (*i.e.* $-\gamma np$). Bimolecular recombination is, as noted above, the reverse process to solar generation since the potential energy lost by the conduction band electron is emitted as a photon. Indeed bimolecular recombination is responsible for light emission in an LED (light emitting diode), which is a device that turns electrical energy into light energy. However other recombination mechanisms are possible and indeed turn out to be more significant that bimolecular recombination in photovoltaic devices. These alternative forms of recombination usually occur when an electron from the conduction band becomes trapped in an energy state lying in the forbidden energy. Such states are termed trap states and are associated with imperfections in the semiconductor. The most common model for recombination via trapped states is the Shockley-Reid-Hall (SRH) model that states that the rate of recombination R_{trap} has the form

$$R_{trap} = \frac{np - n_i^2}{\tau_1 n + \tau_2 p + \tau_3 n_i}, \tag{55}$$

where τ_1, τ_2 and τ_3 are timescales associated with the trapping process. Other more exotic recombination processes are discussed in [6, 9].

4.3 The full equations

We are now in a position to write down the full equations for a semiconductor, which may be doped, and which absorbs light to generate charge carrier pairs. These are given by (46)

[3] Assuming that the recombination and generation rates are independent can lead to nonsensical predictions such as devices that produce more power than they absorb through solar radiation.

and the generalisation of (53) to include solar generation G_{sol} and recombination via trapped states R_{trap}

$$\frac{\partial p}{\partial t} + \frac{1}{q}\nabla \cdot \boldsymbol{j}_p = \gamma(n_i^2 - np) + G_{sol}(\boldsymbol{x}) - R_{trap} \quad \text{where} \quad \boldsymbol{j}_p = M_p p \nabla E_{F_p}, \tag{56}$$

$$\frac{\partial n}{\partial t} - \frac{1}{q}\nabla \cdot \boldsymbol{j}_n = \gamma(n_i^2 - np) + G_{sol}(\boldsymbol{x}) - R_{trap} \quad \text{where} \quad \boldsymbol{j}_n = M_n n \nabla E_{F_n}, \tag{57}$$

$$\nabla \cdot (\varepsilon \nabla \phi) = q(n - p - C_{fix}), \tag{58}$$

where the quasi-Fermi levels are (as in (52)) defined by

$$E_{F_p} = -k_B T \ln\left(\frac{p}{N_v}\right) - (E_{ea} + E_g) - q\phi, \tag{59}$$

$$E_{F_n} = k_B T \ln\left(\frac{n}{N_c}\right) - E_{ea} - q\phi. \tag{60}$$

4.3.1 Boundary conditions.

Boundary conditions on these equations need to be specified where the semiconductor forms a junction with a current collectors. The latter are usually metallic or have metallic properties. In most devices it is usual to assume that the metallic contact and the semiconductor are in thermal equilibrium. Even this assumption still yields rather complicated boundary conditions which typically result in a solution with a thin layer in which the potential changes rapidly. However in practice the contacts are usually chosen so that they do not unduly influence the behaviour of the device and where this is the case a sensible set of boundary conditions, at the interface between a doped semiconductor and a metallic contact, consists of imposing continuity of the potential at the contacts, requiring that the minority carrier current flux across the boundary is zero and ensuring that the carrier charge density is equal to the doping charge density: that is

$$\boldsymbol{J}_{min} \cdot \boldsymbol{n}|_{\partial\Omega} = 0, \quad \phi|_{\partial\Omega} = V, \quad n - p|_{\partial\Omega} = C_{fix}. \tag{61}$$

Here V is the potential of the contact, $\boldsymbol{n}$ is the unit normal to the boundary and $\boldsymbol{J}_{min}$ is the current flux of the minority carrier.

4.4 A simple one-dimensional inorganic solar cell: the n-p homojunction

Here we will consider a solar cell formed from a planar inorganic semiconductor sandwiched between two contacts (see figure 3) at $x = -L$ and $x = L$ and look to calculate its current voltage curve. In order to separate solar generated charges the semiconductor is n-doped in $-L < x < 0$ and p-doped in $0 < x < L$. At equilibrium, as illustrated in figure 4, this results in band bending and a potential difference between the two sides of the cell that acts to drive the electrons and holes created by solar generation apart. In order to simplify the

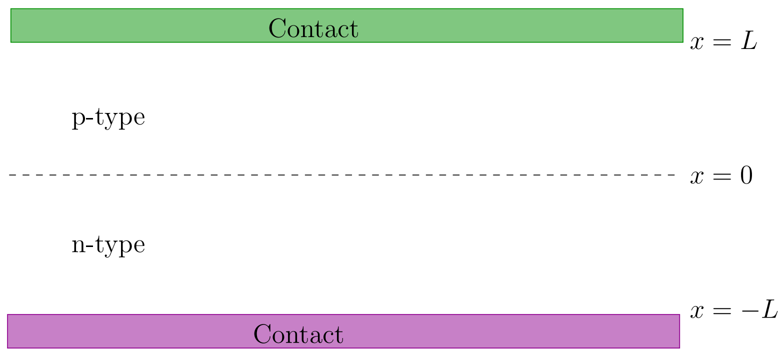

Figure 3: The n-p junction.

problem consider a cell with a high degree of symmetry. In particular assume that the doping in the two halves of the cell is equal and opposite (such that $C_{fix} = N_{dop}$ in $-L < x < 0$ and $C_{fix} = -N_{dop}$ in $0 < x < L$) and that the mobilities of holes and electrons are equal. We shall also assume (rather unrealistically) that recombination is entirely bimolecular (so that $R_{trap} \equiv 0$). With these assumptions the governing equations (56)-(60) may be written as

$$\frac{\partial j_p}{\partial x} = q\left(\gamma(n_i^2 - np) + G_{sol}\right), \qquad j_p = -qD\left(\frac{\partial p}{\partial x} + \frac{p}{V_T}\frac{\partial \phi}{\partial x}\right), \tag{62}$$

$$\frac{\partial j_n}{\partial x} = -q\left(\gamma(n_i^2 - np) + G_{sol}\right), \qquad j_n = qD\left(\frac{\partial n}{\partial x} - \frac{n}{V_T}\frac{\partial \phi}{\partial x}\right), \tag{63}$$

$$\frac{\partial^2 \phi}{\partial x^2} = \begin{cases} \frac{q}{\varepsilon}(n - p - N_{dop}) & \text{in} \quad -L < x < 0 \\ \frac{q}{\varepsilon}(n - p + N_{dop}) & \text{in} \quad 0 < x < L \end{cases}. \tag{64}$$

Here we have chosen to rewrite mobility M in terms of the diffusion coefficient via $D = k_B T M/q$ and have defined $V_T = k_B T/q$. The quantity V_T, termed the thermal voltage, plays an important role in semiconductor physics and at room tempertature $V_T \approx 25$mV.

Boundary Conditions at the contacts. Here we use the boundary conditions of the form (61) at both contacts. On the left contact ($x = -L$) the minority carriers are holes while on the right contact they are electrons. We choose to split voltage difference V between the two contacts into two parts, the built in voltage V_{bi} and the applied voltage V_{ap}, by writing $V = V_{bi} - V_{ap}$; the boundary conditions then become

$$j_n|_{x=-L} = 0, \qquad \phi|_{x=-L} = \frac{V_{bi} - V_{ap}}{2}, \qquad n - p|_{x=-L} = N_{dop}, \tag{65}$$

$$j_p|_{x=L} = 0, \qquad \phi|_{x=L} = \frac{V_{ap} - V_{bi}}{2}, \qquad p - n|_{x=-L} = N_{dop}, \tag{66}$$

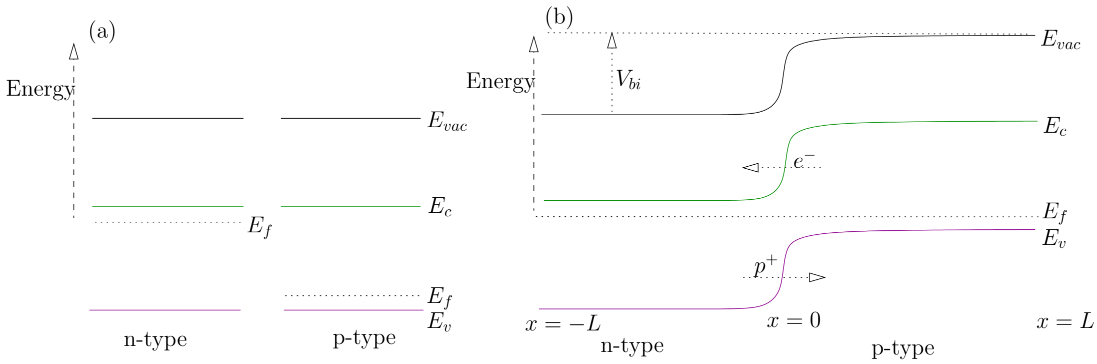

Figure 4: (a) Energy levels in the isolated n-type and p-type materials. (b) Band bending and the built voltage V_{bi} when two materials form an n-p junction.

4.4.1 Rescaling.

We now have the option of trying to solve this model computationally for physically realistic parameters. However this does not really allow us to develop much physical intuition about the problem and it turns out that a much better way of proceding is to try and solve the model asymptotically (*i.e.* approximately) making use of the fact that certain dimensionless parameters (these are sometimes termed dimensionless groups) are either very small or very large in all practical devices. As an example of this let us consider the ratio of the ratio of the Debye length L_D (*i.e.* the characteristic lengthscale defined by equation (64)) to the half-width of the cell. Here on assuming that the characteristic volatge of the problem is V_T and the characteristic carrier density is N_{dop} we find that

$$L_D = \left(\frac{\varepsilon V_T}{q N_{dop}}\right)^{1/2}.$$

Given $\varepsilon \sim 4\varepsilon_0$ and $N_{dop} \sim 10^{22}-10^{25}\text{m}^{-3}$ gives L_D in the range $10^{-9}-10^{-8}$m. Comparing this to a typical half width of a cell, which might be $10^{-6}-10^{-5}$m, implies that the dimensionless parameter λ defined by

$$\lambda = \frac{L_D}{L} = \frac{1}{L}\left(\frac{\varepsilon V_T}{q N_{dop}}\right)^{1/2} \tag{67}$$

is very small $\lambda = O(10^{-4}) - O(10^{-2})$, a fact that we shall make use of.

In order to identify all the dimensionless parameters in the problem we rescale it in such a way as to leave a dimensionless problem charactersied by a minimal set of dimensionless parameters. This process is called non-dimensionalisation and, providing that we choose to scale our variables with sensible values, leads to considerable insight into the physics of the problem.

Nondimensionalisation. We resacle the variables in the problem as follows:

$$\phi = V_T\bar{\phi}, \quad n = N_{dop}\bar{n}, \quad p = N_{dop}\bar{p}, \quad G_{sol} = G_{1-sun}\bar{G}, \quad j_n = qLG_{1-sun}\bar{j}_n,$$
$$x = L\bar{x}, \quad V_{bi} = V_T\bar{V}_{bi}, \quad V_{ap} = V_T\bar{V}_{ap} \quad j_p = qLG_{1-sun}\bar{j}_p.$$

where given the application we choose to scale the charge pair generation term with the generation that would be expected under illumination of one sun G_{1-sun}. On substituting these rescalings into (62)-(66) we obtain the following system on dimensionless equations and boundary conditions:

$$\frac{\partial \bar{j}_p}{\partial \bar{x}} = \Theta(N_i^2 - \bar{n}\bar{p}) + \bar{G}, \qquad \bar{j}_p = -\kappa\left(\frac{\partial \bar{p}}{\partial \bar{x}} + \bar{p}\frac{\partial \bar{\phi}}{\partial \bar{x}}\right), \tag{68}$$

$$\frac{\partial \bar{j}_n}{\partial \bar{x}} = -\Theta(N_i^2 - \bar{n}\bar{p}) - \bar{G}, \qquad \bar{j}_p = \kappa\left(\frac{\partial \bar{n}}{\partial \bar{x}} - \bar{n}\frac{\partial \bar{\phi}}{\partial \bar{x}}\right), \tag{69}$$

$$\frac{\partial^2 \bar{\phi}}{\partial \bar{x}^2} = \begin{cases} \frac{1}{\lambda^2}(\bar{n} - \bar{p} - 1) & \text{in} \quad -1 < \bar{x} < 0 \\ \frac{1}{\lambda^2}(\bar{n} - \bar{p} + 1) & \text{in} \quad 0 < \bar{x} < 1 \end{cases}, \tag{70}$$

$$\bar{j}_p|_{\bar{x}=-1} = 0, \qquad \bar{\phi}|_{\bar{x}=-1} = \frac{\bar{V}_{bi} - \bar{V}_{ap}}{2}, \qquad \bar{n} - \bar{p}|_{\bar{x}=-1} = 1, \tag{71}$$

$$\bar{j}_n|_{\bar{x}=1} = 0, \qquad \bar{\phi}|_{\bar{x}=1} = \frac{\bar{V}_{ap} - \bar{V}_{bi}}{2}, \qquad \bar{p} - \bar{n}|_{\bar{x}=1} = 1, \tag{72}$$

where the dimensionless parameters (or groups) in the problem are

$$N_i = \frac{n_i}{N_{dop}}, \quad \lambda = \frac{1}{L}\left(\frac{\varepsilon V_T}{qN_{dop}}\right)^{1/2}, \quad \Theta = \left(\frac{\gamma N_{dop}^2}{G_{1-sun}}\right), \quad \kappa = \left(\frac{DN_{dop}}{G_{1-sun}L^2}\right). \tag{73}$$

Dimensionless parameters and their size. We have already seen that $\lambda = L_D/L$ gives the ratio of the Debye length to the half-width of the cell. The next parameter we consider is N_i which gives the ratio of the intrinsic carrier density to that of the doping density; the intrinsic carrier density in silicon (at room temperature) is roughly 10^{16}m^{-3} and so taking $N_{dop} \sim 10^{20} - 10^{22}\text{m}^{-3}$ (as above) leads to an estimate of $N_i = O(10^{-4}) - O(10^{-8})$. Θ gives the ratio of the square of the typical charge densities obtained by balancing solar generation with bimolecular recombination to the square of the doping density. κ gives the ratio of the typical current density produced by a potential difference across the cell of size $O(V_T)$ to the current density produced by charge generation at one sun. Thus if $\kappa \gg 1$ this suggests the current produced by illumination of the solar cell at one sun can be transported away by a very small potential diffference across the cell, or to put it another way that the ohmic resistance is negligible. It is nearly always the case that $\kappa \gg 1$, indeed if it were not the cell would be very inefficient.

Simplifying the problem using its symmetry. The problem is symmetric about $\bar{x} = 0$ as can be seen by making the substitution

$$\bar{x} = -\bar{X}, \quad \bar{p} = \bar{N}, \quad \bar{n} = \bar{P}, \quad \bar{\phi} = -\bar{\Phi}, \quad \bar{j}_n = \bar{J}_p, \quad \bar{j}_p = \bar{J}_n,$$

in (68)-(72). If you do this you will see that you retrieve exactly the same problem but with the variables replaced as follows:

$$\bar{x} \to \bar{X}, \quad \bar{n} \to \bar{N}, \quad \bar{p} \to \bar{P}, \quad \bar{\phi} \to \bar{\Phi}, \quad \bar{j}_n \to \bar{J}_n, \quad \bar{j}_p \to \bar{J}_p,$$

(check this). It is thus apparent that the solution to (68)-(72) has the symmetry

$$\bar{p}(-\bar{x}) = \bar{n}(\bar{x}), \quad \bar{n}(-\bar{x}) = \bar{p}(\bar{x}), \quad \bar{\phi}(-\bar{x}) = -\bar{\phi}(\bar{x}), \tag{74}$$

$$\bar{j}_p(-\bar{x}) = \bar{j}_n(\bar{x}), \quad \bar{j}_n(-\bar{x}) = \bar{j}_p(\bar{x}). \tag{75}$$

We need therefore only solve in $0 < \bar{x} < 1$ where we impose the boundary conditions

$$\bar{p}|_{\bar{x}=0} = \bar{n}|_{\bar{x}=0}, \quad \bar{\phi}|_{\bar{x}=0} = 0, \quad \bar{j}_p|_{\bar{x}=0} = \bar{j}_n|_{\bar{x}=0}.$$

The problem we need to solve in $0 < \bar{x} < 1$ can thus be stated as

$$\frac{\partial \bar{j}_p}{\partial \bar{x}} = \Theta(N_i^2 - \bar{n}\bar{p}) + \bar{G}, \qquad \bar{j}_p = -\kappa \left(\frac{\partial \bar{p}}{\partial \bar{x}} + \bar{p} \frac{\partial \bar{\phi}}{\partial \bar{x}} \right), \tag{76}$$

$$\frac{\partial \bar{j}_n}{\partial \bar{x}} = -\Theta(N_i^2 - \bar{n}\bar{p}) - \bar{G}, \qquad \bar{j}_p = \kappa \left(\frac{\partial \bar{n}}{\partial \bar{x}} - \bar{n} \frac{\partial \bar{\phi}}{\partial \bar{x}} \right), \tag{77}$$

$$\frac{\partial^2 \bar{\phi}}{\partial \bar{x}^2} = \frac{1}{\lambda^2}(\bar{n} - \bar{p} + 1), \tag{78}$$

$$\bar{p}|_{\bar{x}=0} = \bar{n}|_{\bar{x}=0}, \qquad \bar{\phi}|_{\bar{x}=0} = 0, \qquad \bar{j}_p|_{\bar{x}=0} = \bar{j}_n|_{\bar{x}=0},, \tag{79}$$

$$\bar{p} - \bar{n}|_{\bar{x}=1} = 1, \qquad \bar{\phi}|_{\bar{x}=1} = \frac{\bar{V}_{ap} - \bar{V}_{bi}}{2}, \qquad \bar{j}_n|_{\bar{x}=1} = 0, \tag{80}$$

and we can determine the solution in $-1 < \bar{x} < 0$ from the symmetry conditions (75).

4.4.2 Asymptotic solution to the dimensionless model in the limit $\lambda \ll 1$ and $\kappa \gg 1$

We now seek to solve the problem approximately by making use of the fact that $\kappa \gg 1$ and $\lambda \ll 1$. We start by noting that since $\kappa \gg 1$ the right-hand sides of (76) and (77) must both be very small, *i.e.* the following approximate equations must hold

$$\frac{\partial \bar{p}}{\partial \bar{x}} \sim -\bar{p} \frac{\partial \bar{\phi}}{\partial \bar{x}}, \qquad \frac{\partial \bar{n}}{\partial \bar{x}} \sim \bar{n} \frac{\partial \bar{\phi}}{\partial \bar{x}}.$$

It is not hard to show that these have solution

$$\bar{p} = A \exp(-\bar{\phi}) \quad \text{and} \quad \bar{n} = B \exp(\bar{\phi}), \tag{81}$$

for some arbitrary constants A and B. Apllying the conditions (79a-b) on $\bar{x} = 0$ implies that

$$B = A$$

while application of the conditions (80a-b) on $\bar{x} = 1$ implies that

$$A = \frac{1}{\exp\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right) - \exp\left(-\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)} = \frac{1}{2\sinh\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)},$$

so that (81) can now be rewritten

$$\bar{p} = \frac{\exp(-\bar{\phi})}{2\sinh\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)}, \qquad \bar{n} = \frac{\exp(\bar{\phi})}{2\sinh\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)}. \tag{82}$$

Substituting for $\bar{p}$ and $\bar{n}$ in (78) we find the following Poisson-Boltzmann equation for $\bar{\phi}$:

$$\frac{\partial^2 \bar{\phi}}{\partial \bar{x}^2} = \frac{1}{\lambda^2}\left(\frac{\sinh(\bar{\phi})}{\sinh\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)} + 1\right), \tag{83}$$

with boundary conditions

$$\bar{\phi}|_{\bar{x}=0} = 0 \quad \text{and} \quad \bar{\phi}|_{\bar{x}=1} = -\frac{\bar{V}_{bi} - \bar{V}_{ap}}{2}. \tag{84}$$

The small λ limit. Now we know that $\lambda \ll 1$. This means that the term in the brackets on the right-hand side of (83) must be very small, *i.e.*

$$\frac{\sinh(\bar{\phi})}{\sinh\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)} \sim -1.$$

The only solution to this approximate equation is

$$\bar{\phi} \sim -\frac{\bar{V}_{bi} - \bar{V}_{ap}}{2}, \tag{85}$$

which obviously satisfies the second of the boundary conditions (84) but not the first. The problem here is that we cannot expect to neglect the highest derivative in an ODE and still satisfy its boundary conditions so that our naive approximation (85) fails. Problems in which we are tempted to neglect the highest derivative, but get into trouble doing so, are called singular perturbation problems. And, as we shall see, their solutions often have boundary layers in which the solution varies rapidly. Indeed, as we shall show, this is precisely what happens here and in fact

$$\bar{\phi} \sim -\frac{\bar{V}_{bi} - \bar{V}_{ap}}{2} \qquad \text{for} \quad \lambda \ll x < 1. \tag{86}$$

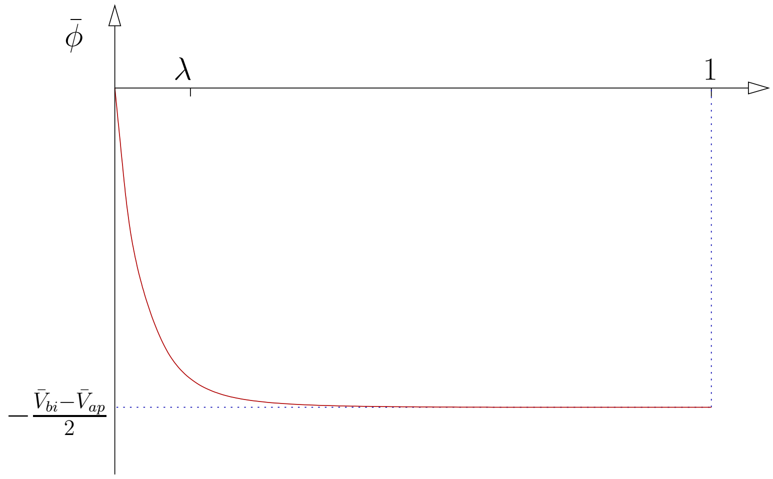

Figure 5: Illustration of the solution of (83)-(84) for $\bar{\phi}$ with $\lambda \ll 1$.

That is the approximation only fails in a narrow region of size $O(\lambda)$ around $\bar{x} = 0$ (this is illustrated in figure 5). If the solution is indeed as we have postulated in figure 5 then we ought to be able to investigate the rapid transition from $\bar{\phi} = 0$ at $\bar{x} = 0$ to the approximate solution (86) for $\bar{x} \gg \lambda$ by rescaling $\bar{x}$ with λ by writing

$$\bar{x} = \lambda z$$

On making this substitution into (83) and (84) we get

$$\frac{\partial^2 \bar{\phi}}{\partial z^2} = \left(\frac{\sinh(\bar{\phi})}{\sinh\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)} + 1 \right) \quad \text{with} \quad \bar{\phi}|_{z=0} = 0, \tag{87}$$

and if it is indeed to tend towards the approximate solution (86), for large z, we must impose the far-field condition

$$\bar{\phi} \to -\frac{\bar{V}_{bi} - \bar{V}_{ap}}{2} \quad \text{as} \quad z \to +\infty. \tag{88}$$

Provided you believe that there is a solution to (87) satisfying the far-field condition (88) it is now fairly easy to find a good approximation to the current voltage curve for the device. In fact it turns out that there is a relatively straightforward to show that such a solution exists provided that $\bar{V}_{bi} - \bar{V}_{ap} > 0$ by using a phase-plane analysis.

Excercise: Write $w = \bar{\phi}_z$ and then write the ODE (83a) as two coupled first order autonomous ODEs of the form $\bar{\phi}_z = w$ and and $w_z = f(\bar{\phi})$. Sketch the phase-plane for these coupled ODEs and show that there is a trajectory linking $\bar{\phi} = 0$ to a critical point at $w = 0$, $\bar{\phi} = -(\bar{V}_{bi} - \bar{V}_{ap})/2$ when $\bar{V}_{bi} - \bar{V}_{ap} > 0$ and hence argue that a solution to the problem (87)-(88) exists when $\bar{V}_{bi} - \bar{V}_{ap} > 0$.

Calculating the current voltage curves (I). Now we know $\bar{n}$ and $\bar{p}$ nearly everywhere in $0 < \bar{x} < 1$ except very close to $\bar{x} = 0$. It follows that we can get a very good approximation to the amount of recombination occuring and hence work out a current voltage curve. Since $\bar{\phi} \sim (\bar{V}_{ap} - \bar{V}_{bi})/2$ in almost all of $0 < \bar{x} < 1$ it follows from (82) that

$$\bar{n} \sim \frac{\exp\left(-\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)}{2\sinh\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)}, \qquad \bar{p} \sim \frac{\exp\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)}{2\sinh\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)}, \tag{89}$$

The built-in voltage $\bar{V}_{bi}$. Lets look first at what happens when there is zero solar generation $\bar{G} \equiv 0$ and zero applied voltage $\bar{V}_{ap} = 0$ and the device is allowed to reach equilibrium so that $\bar{j}_n = \bar{j}_p = 0$. These assumptions combined with (76a) and (77a) imply that $\bar{n}\bar{p} = N_i^2$. On subsituting the values of $\bar{n}$ and $\bar{p}$, as determined in (89) with $\bar{V}_{ap} = 0$, into this relation gives an equation for $\bar{V}_{bi}$

$$\sinh^2\left(\frac{\bar{V}_{bi}}{2}\right) = \frac{1}{4N_i^2}.$$

This can be further simplified by noting that (in all practical applications) N_i the ratio of the intrinsic carrier density to the doping density is very small which implied that $\bar{V}_{bi}$ is large and given (approximately) by

$$\bar{V}_{bi} \sim 2\ln\left(\frac{1}{N_i}\right). \tag{90}$$

Translating this back into dimensional variables (via (68)) this gives

$$V_{bi} = 2V_T \ln\left(\frac{N_{dop}}{n_i}\right), \tag{91}$$

which is precisely the result we would get from looking at the band bending and lining up the Fermi levels (which is reassuring).

Calculating the current voltage curves (II). We now look at the case where the generation term $\bar{G}$ (which is assumed constant) and the applied voltage $\bar{V}_{ap}$ are both non-zero and attempt to calculate the total current density $\bar{j} = \bar{j}_n + \bar{j}_p$ as a function of the applied voltage. The equations and boundary conditions for the current densities become, on substituting for $\bar{n}$ and $\bar{p}$ from (89) in (76a), (77a), (79b) and (80a),

$$\frac{\partial \bar{j}_p}{\partial \bar{x}} = \Theta\left(N_i^2 - \frac{1}{4\sinh^2\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)}\right) + \bar{G}, \tag{92}$$

$$\frac{\partial \bar{j}_n}{\partial \bar{x}} = -\Theta\left(N_i^2 - \frac{1}{4\sinh^2\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)}\right) - \bar{G}, \tag{93}$$

$$\bar{j}_p|_{\bar{x}=0} = \bar{j}_n|_{\bar{x}=0}, \qquad \bar{j}_n|_{\bar{x}=1} = 0. \tag{94}$$

Notably the terms on the right-hand side of (92) and (93) are constant so that they are strightforward to integrate. In particular

$$[\bar{j}_p - \bar{j}_n]_{\bar{x}=0}^{1} = 2\left(\bar{G} - \Theta\left(\frac{1}{4\sinh^2\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)} - N_i^2\right)\right)$$

so that, since $\bar{j}_p - \bar{j}_n|_{\bar{x}=0} = 0$ (from (79c)) and $\bar{j}_n|_{\bar{x}=1} = 0$ (from (80c)), it follows that the dimensionless current density $\bar{j} = \bar{j}_n = \bar{j}_p$ is given by

$$\bar{j} = \bar{j}_p|_{\bar{x}=1} = 2\left(\bar{G} - \Theta\left(\frac{1}{4\sinh^2\left(\frac{\bar{V}_{bi}-\bar{V}_{ap}}{2}\right)} - N_i^2\right)\right).$$

Furthermore since $\bar{V}_{bi} - \bar{V}_{ap}$ is typically large (*i.e.* in dimensional terms many thermal voltages) we can approximate $4\sinh^2\left((\bar{V}_{bi} - \bar{V}_{ap})/2\right) \sim \exp(\bar{V}_{bi} - \bar{V}_{ap})$ so that, on substituting for V_{bi} from (90), we obtain

$$\bar{j} \sim 2\left(\bar{G} - N_i^2\Theta(e^{\bar{V}_{ap}} - 1)\right). \tag{95}$$

Rewriting this in terms of the original dimensional variables, using (68) and (73), we find

$$j = 2qL\left(G - \gamma n_i^2\left(\exp\left(\frac{V_{ap}}{V_T}\right) - 1\right)\right). \tag{96}$$

As we shall show in the §5 this corresponds precisely to the current voltage relation of a current source strenght $2qLG$ in parallel with an ideal diode (ideality factor 1) with reverse saturation current density $2qL\gamma n_i^2$.

Excercise: Modify the above argument to calculate the current voltage curve where recombination primarily occurs via traps and can be modelled by a Shockley-Reid-Hall recombination term as given in (55).

5 Shockley equivalent circuit models of photovoltaic devices

Drift diffusion models of charge transport, although they are capable of accurately describing the behaviour of photovoltaics, have the disadvantage of being complicated and hard to solve. Indeed numerical solutions of these models in the right parameter regimes are frequently very difficult to obtain with any degree of accuracy. When comparing to experimental data it is often easier to make use of phenomenological equivalent circuit models. At their very simplest these consist of a current source term (representing the current generated by solar radiation) in parallel with a diode (see figure 6(a)) which represents the effects of charge recombination. Indeed, as we shall show, the current-voltage curve that we obtained from our approximate analysis of a drift-diffusion model of an np-homojunction (96) is precisely of this form.

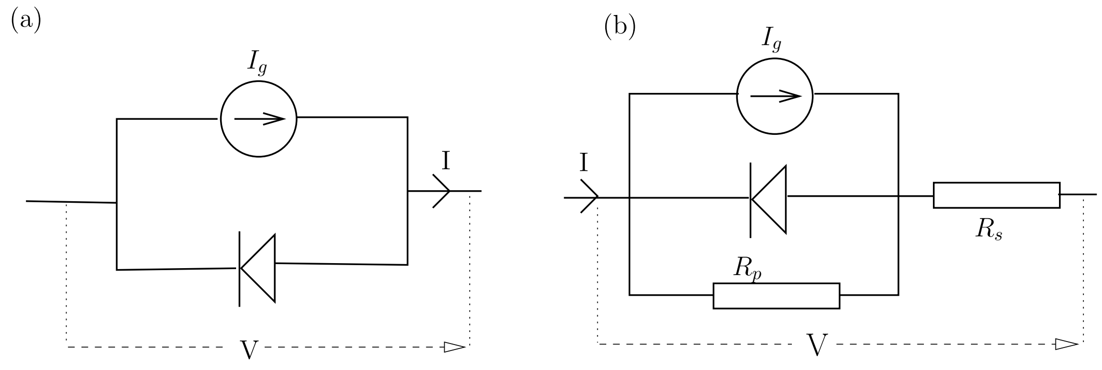

Figure 6: Shockley equivalent circuits for a solar cell. (a) simple circuit for an n-p diode (b) the generic Shockley equivalent circuit.

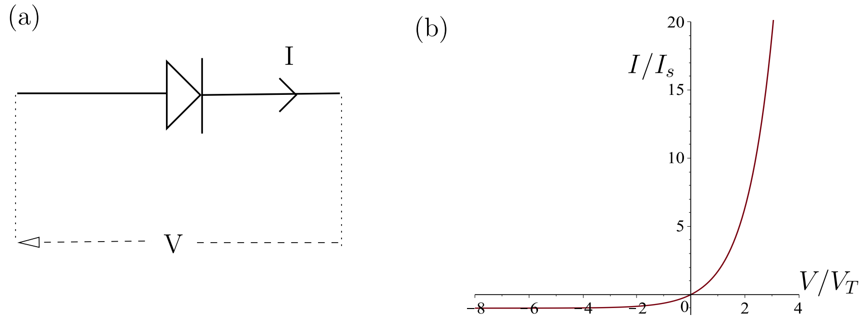

Figure 7: (a) Illustration of the direction of the current flow I through, and voltage V across, a diode. (b) The current-voltage curve for a diode with ideality $\mathcal{N}_{id} = 1$.

The current voltage curve of a diode. The current I through a diode is related to the voltage across it V by the relation

$$I = I_s \left[\exp\left(\frac{V}{\mathcal{N}_{id} V_T} \right) - 1 \right], \tag{97}$$

where I_s and $\mathcal{N}_{id}$ are termed the reverse saturation current and the ideality factor respectively. Since this relation is asymmetric the diode (as a device) has directionality and this is illustrated in figure 7.

The simple Shockley equivalent circuit (figure 6a). We now use the diode current-voltage curve, defined in (97), in the simple equivalent circuit in figure 5a. Note that the diode figure 5a points in the opposite direction to that in figure 7. The current flowing through the diode, from left to right, is thus

$$I_{diode} = -I_s \left[\exp\left(\frac{V}{\mathcal{N}_{id} V_T} \right) - 1 \right].$$

addiing this current to the current I_g from the current source gives the total current I flowing through the circuit

$$I = I_g - I_s \left[\exp\left(\frac{V}{\mathcal{N}_{id} V_T} \right) - 1 \right]. \tag{98}$$

Note that this current-voltage curve has exactly the same form as the current-voltage curve we predicted from the n-p homojunction in §4.4, *i.e.* (96), provided we identify

$$I_g = 2AqLG, \quad I_s = 2Aq\gamma n_i^2, \quad \mathcal{N}_{id} = 1.$$

Here A is the area of the homojunction so that $I = jA$. One rather interesting consequence of assuming that recombination is entirely bimolecular is the prediction that the ideality factor $\mathcal{N}_{id} = 1$. If instead we had assumed that the dominant form of recombination were SRH recombination then we would have found that the ideality factor $\mathcal{N}_{id} = 2$. Other forms of recombination give other ideality factors, for example Auger recombination (see [7]) gives an ideality factor of 3.

Other devices. Whilst our analysis of the n-p homojunction in §4.4 shows that it is appropriate to fit current voltage data for such a device to an equivalent circuit model it is not obvious that this approach is appropriate for other devices. Roughly speaking if the device has planar material and doping interfaces that run parallel to the contacts an equivalent ciruit description will provide a consistent approximation of its behaviour. In [2] an analysis of a planar organic bulk heterojunction is made which results in a current-voltage curve that is well approximated by a Shockley equivalent circuit whilst in [3] a similar type analysis is conducted for a planar (tri-layer) perovskite cell. It should be emphasised that cells with non-planar geometry cannot usually be modelled by a Shockley equivalent circuit (*e.g.* heterogeneous organic bulk heterojunctions).

The generic Shockley equivalent circuit (figure 6b). The generic Shockley equivalent circuit contains a shunt resistance R_p in parallel with a diode and a current source. The three parallel elements are connected in series with a resistor R_s. By splitting the voltage into two parts: V_1 (the voltage drop acoss diode, shunt resistance and current source) and V_2 the voltage drop across R_s, we obtain the following set of equations

$$\begin{aligned}
&V = V_1 + V_2, && \\
&I_{diode} = -I_s \left[\exp\left(\frac{V_1}{\mathcal{N}_{id} V_T} \right) - 1 \right] && \text{Current through diode,} \\
&I_{shunt} = -\frac{V_1}{R_p} && \text{Current through shunt resistor,} \\
&I = I_g + I_{shunt} + I_{diode} && \text{Conservation of current (Kirchoff),} \\
&V_2 = -IR_s && \text{Current through series resistor.}
\end{aligned}$$

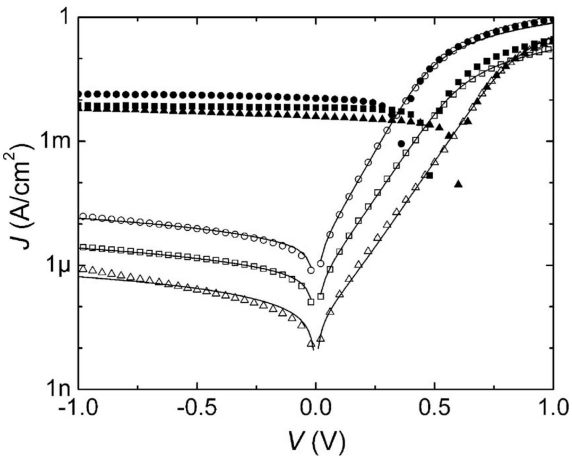

Figure 8: Fitting an equivalent circuit model to current-voltage data from bilayer heterojunction organic solar cells (see [8]). Filled symbols represent cells under illumination open symbols cells in the dark. Notice the current has been plotted on a log scale.

It is straightfoward to combine these expressions to obtain a single implicit expression for the current voltage curve

$$I = I_g - I_s \left[\exp\left(\frac{V + IR_s}{\mathcal{N}_{id} V_T} \right) - 1 \right] - \frac{V + IR_s}{R_p}. \tag{99}$$

What is the physical basis for series and shunt resistances. We have already seen that charge recombination in an n-p homojunction leads to the diode like behaviour of this device. So what processes give rise to the shunt resistance and the series resistance used to model real photovoltaic devices? The series resistance R_s can arise as a result of the resistance of the contacts or because of the resistance of the semiconductor itself. The latter effect can be fairly readily incorporated into the analysis that we conducted in §4.4. The shunt resistance is in some sense more interesting because it usually arises from shorting between the contacts. This might arise because there are regions where one of the doped layers are absent and the other spans the entire width of the cell or because dendrites have grown directly from one contact to the other. In either case a low value of R_p can be attributed to some fault in the manufacture of the cell. This is therefore clearly something to look out for when fitting current voltage curves to an equivalent circuit.

Use of equivalent circuits models to fit to real data. Here we show a couple of examples of fits to real current-voltage data. In figure 8 we show fitted data taken from Potscavage et al. [8] for current-voltage curves measured from three different designs of bilayer heterojunction organic solar cells. In the cases of the data plotted with triangles and squares the ideality factor $\mathcal{N}_{id} \approx 2$ corresponding to SRH recombination, while in the case of data plotted with circles $\mathcal{N}_{id} \approx 1.7$ In figure 9 we show a comparison of an experimental current voltage curve obtained from a perovskite cell and the fit using a simple equivalent

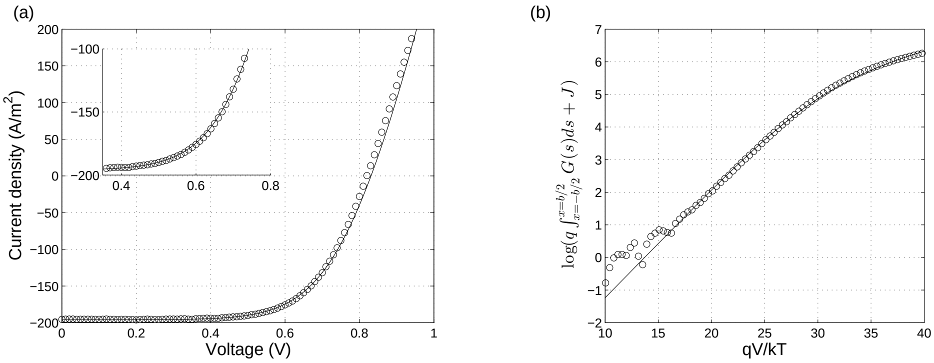

Figure 9: Fitting an equivalent circuit model to current-voltage data from a perovskite cell (prepared by Giles Eperon). From the data it is inferred that the ideality factor $\mathcal{N}_{id} \approx 3$. For more details see [3]. Note that the current has been defined in an opposite sense to that in the text (to agree with the text the figure should be flipped about the V-axis).

circuit (taken from [3]). Here the ideality factor is $\mathcal{N}_{id} \approx 3$.Whilst this high ideality factor might be attributable to Auger recombination we believe it is more likely related to the hysteretic behaviour of perovskite cells.

References

[1] R.P. Feynman, R.B. Leighton & M. Sands, "Lectures on Physics (Volume III)". Addison Wesley (2003).

[2] J. M. Foster, J. Kirkpatrick & G. Richardson, "Asymptotic and numerical prediction of current-voltage curves for an organic bilayer solar cell under varying illumination and comparison to the Shockley equivalent circuit". J. Appl. Phys., **114**: 104501 (2013).

[3] Foster J.M., Snaith H.J., Leijtens T. & Richardson G., "A model for the operation of perovskite based hybrid solar cells: formulation, analysis and comparison to experiment". *SIAM J. Appl. Math.* **74**, 1935-1966 (2014).

[4] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions". *J. Phys. Chem.* **81**, 2340-2361 (1977).

[5] Guarnera S., Abate A., Zhang W., Foster J.M., Richardson G., Petrozza A. & Snaith H.J. "Improving the long-term stability of perovskite solar cells with a porous Al_2O_3 overlayer". *J. Phys. Chem. Lett.* **6**, 432-437 (2015).

[6] T. Kirchartz, B. E. Pieters, J. Kirkpatrick, U. Rau & J. Nelson, "Recombination of tail states in polythiopene: fullerene solar cells". Phys. Rev. B., **83**:115209 (2011).

[7] J. Nelson, "The physics of Solar cells". Imperial College Press, (2003).

[8] W. J. Potscavage, S. Yoo & B. Kippelen, 'Origin of the open-circuit voltage in a multillayer heterojunction organic solar cell', Appl. Phys. Lett. **93**, 193308 (2008).

[9] A. Schenk & U. Krumbein, "Coupled defect-level recombination: Theory and application to anomolous diode characteristics". J. Appl. Phys., **78**: 3185-3192 (1995).

[10] C. Schmeiser & A. Unterreiter, "The derivation of analytic device models by asymptotic methods". Semiconductors, Part II, IMA Vol. in Math. and its Appl. 59, Springer-Verlag (1993).

[11] S. M. Sze, K. Ng Kwok, "Physics of Semiconductor Devices". 3rd edition, Wiley-Interscience NY, (2006).

[12] A. B. Walker, "Monte Carlo studies of electronic processes in dye-sensitized solar cells", Topics in Current Chemistry (2013).