PRINCIPLES OF A-SI:H PHOTODIODES FOR MEDICAL X-RAY IMAGING

by

Bernard McAuley B.Sc., M.Sc.

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy with the Department of Physics in the University of Surrey

1998

© Bernard McAuley, September 1998
Errata

Pg. 4 X-rays were discovered in December 1895 (not 1886 as given in the text).

Pg. 11, Fig 1-4 CsI thickness is 100 μm not 100 mm.

Pg. 33, Fig 2-2 label on figure should read “Donor and Acceptor states” not just “Donor states”.

Pg. 55, Eqn 2-45 Final term “dE” should be “dE_i”.

Pg. 69 second paragraph last line should read “that this is in fact” instead of “that this in fact”.

Pg. 92/93, Fig 3-5 and 3-6 figure legends are incorrectly labelled ID#1 and should in fact be ID#2.
Abstract

In this thesis we assess the use of a-Si photodiodes for medical X-ray imaging. The design of an a-Si X-ray imager based on an active matrix, suitable for medical imaging is outlined. The key points in the design are highlighted, and the role of the a-Si photodiode is discussed. We outline the underlying physics of photodiodes, with particular attention to a-Si, and review the current literature. Based on this a model suitable for simulation has been developed. This model is able to replicate the thickness independence of a-Si diodes without resorting to an arbitrary charge at the p/i interface.

Because individual pixels must be isolated, a certain fraction of any pixel will have to be contactless. This causes a considerable loss of signal. The possibility of retrieving this lost signal by using a lateral field was examined using the above model. Five structures were examined.

1. A basic photodiode with a longer top contact than the bottom contact.
2. A junction focussed photodiode where an addition junction contact is placed between the contacts on the bottom to induce a lateral field.
3. An insulated focussed photodiode where an insulated contact is placed between the contacts on the bottom of the diode to induce a lateral field.
4. A skewed photodiode, where the top and bottom contacts are deliberately misaligned to generate a lateral potential.
5. A skewed photodiode with additional insulating contacts on the top and bottom to improve the lateral potential.

The focussing elements in these devices were show to have varying degrees of efficiency.

It was noted that the current models for recombination would not be suitable for transient simulations. Also, as they neglect the presence of amphoteric defects it seems likely that they would underestimate the recombination. An examination of the current dangling bond models was undertaken with a view to using these in simulations. However, application of these models to the problem at hand revealed limitations in their use. A new approach was tried, using a transient analysis, but this was also found to be wanting.

Finally, possible paths for future work are discussed, along with the apparatus constructed for measurements to corroborate the simulation results.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ABSTRACT</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>TABLE OF CONTENTS</strong></td>
<td>II</td>
</tr>
<tr>
<td><strong>ACKNOWLEDGEMENTS</strong></td>
<td>V</td>
</tr>
<tr>
<td><strong>CHAPTER 1 ELEMENTS OF DESIGN FOR A DIGITAL MEDICAL X-RAY IMAGER</strong></td>
<td>1</td>
</tr>
<tr>
<td>SECTION 1.1 - INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 Outline</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2 Amorphous Silicon and Large Area Electronics</td>
<td>2</td>
</tr>
<tr>
<td>SECTION 1.2 - MEDICAL X-RAY IMAGING</td>
<td>4</td>
</tr>
<tr>
<td>SECTION 1.3 - DETECTOR DESIGN AND PRINCIPLES</td>
<td>13</td>
</tr>
<tr>
<td>SECTION 1.4 - ADVANTAGES AND DISADVANTAGES OF THE INTEGRATING PIXEL DETECTOR</td>
<td>15</td>
</tr>
<tr>
<td>SECTION 1.5 - REQUIREMENTS FOR PHOTODIODE DETECTORS</td>
<td>16</td>
</tr>
<tr>
<td>SECTION 1.6 - SUMMARY</td>
<td>22</td>
</tr>
<tr>
<td>SECTION 1.7 - REFERENCES</td>
<td>22</td>
</tr>
<tr>
<td><strong>CHAPTER 2 THE PHYSICS OF A-Si:H PHOTODIODES</strong></td>
<td>24</td>
</tr>
<tr>
<td>SECTION 2.1 - INTRODUCTION</td>
<td>24</td>
</tr>
<tr>
<td>SECTION 2.2 - THE PRODUCTION OF A-Si DEVICES</td>
<td>24</td>
</tr>
<tr>
<td>SECTION 2.3 - FROM C TO A: HOW AMORPHOUS MATERIALS DIFFER FROM CRISTALLINE MATERIALS</td>
<td>29</td>
</tr>
<tr>
<td>2.3.1 Structure</td>
<td>29</td>
</tr>
<tr>
<td>2.3.2 Band Theory of Semiconductors and the Density of States (DOS)</td>
<td>31</td>
</tr>
<tr>
<td>2.3.3 Doping of a-Si devices</td>
<td>34</td>
</tr>
<tr>
<td>SECTION 2.4 - WHEN OPPOSITES COLLIDE: THE SEMICONDUCTOR JUNCTION</td>
<td>37</td>
</tr>
<tr>
<td>2.4.1 Introduction</td>
<td>37</td>
</tr>
<tr>
<td>2.4.2 The Basic Semiconductor Equations</td>
<td>38</td>
</tr>
<tr>
<td>2.4.3 The Silicon p-n Diode and the Depletion Region</td>
<td>40</td>
</tr>
<tr>
<td>2.4.4 The Shockley Equation (The ideal diode law)</td>
<td>42</td>
</tr>
<tr>
<td>2.4.5 Non-Ideal Diodes</td>
<td>46</td>
</tr>
<tr>
<td>2.4.6 The p-i-n Diode</td>
<td>50</td>
</tr>
<tr>
<td>SECTION 2.5 - THE A-Si DIODE</td>
<td>52</td>
</tr>
<tr>
<td>2.5.1 Modifications to the Semiconductor equations</td>
<td>52</td>
</tr>
<tr>
<td>2.5.2 Recombination and Generation in a-Si</td>
<td>54</td>
</tr>
<tr>
<td>2.5.3 The a-Si Density of States</td>
<td>60</td>
</tr>
<tr>
<td>2.5.4 Summary</td>
<td>63</td>
</tr>
<tr>
<td>SECTION 2.6 - PREVIOUS WORK ON A-Si DIODES</td>
<td>64</td>
</tr>
<tr>
<td>SECTION 2.7 - THE ELECTRONICS OF PHOTODETECTORS</td>
<td>70</td>
</tr>
<tr>
<td>SECTION 2.8 - SUMMARY</td>
<td>74</td>
</tr>
<tr>
<td>SECTION 2.9 - REFERENCES</td>
<td>75</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.3.3 Software for Controlling the Electro-Optic Probe Station</td>
<td>188</td>
</tr>
<tr>
<td>5.3.4 Some Results Obtained from the Electro-Optic Probe Station</td>
<td>192</td>
</tr>
<tr>
<td>SECTION 5.4 - CONCLUSIONS</td>
<td>194</td>
</tr>
<tr>
<td>SECTION 5.5 - REFERENCES</td>
<td>197</td>
</tr>
<tr>
<td>APPENDIX A AN INTRODUCTION TO NUMERICAL METHODS</td>
<td>A-1</td>
</tr>
<tr>
<td>APPENDIX B DERIVATION OF THE DEFECT POOL MODEL AND THE OCCUPATION</td>
<td>B-1</td>
</tr>
<tr>
<td>APPENDIX C CAPTURE AND EMISSION COEFFICIENTS IN DANGLING BOND RECOMBINATION</td>
<td>C-1</td>
</tr>
<tr>
<td>APPENDIX D SILVACO SOFTWARE</td>
<td>D-1</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I wish to acknowledge Dr E.J. Morton for his supervision during the past three years of research and his input on all aspects of this work. Thanks are also due to Dr W. Gilboy for his help. I would also like to acknowledge Dr M.J. Rose and the other staff in the Amorphous Material Research Group at Dundee University for their help in this project.

Most importantly thanks are due to Serrita, my wife, for her moral support, constant interest and several weeks worth of printing and proof reading without which this thesis would never have seen the light of day.

Honourable mentions are due to my fellow students, without whom coffee breaks would have been much duller and all of whom have helped move this project forward. Special thanks are due to Paul De Antonis whose electrical and mechanical design skills have proven invaluable. Thanks also to Frank Podd, Luis Tavora, Tiago de Menezes, Martyn Key, Ashoke Kundu and Tony Ede who have all contributed to this work. Finally, thanks to Dave Munro for keeping the computers up and running.
Chapter 1  Elements of Design for a Digital Medical X-ray Imager

Section 1.1 - Introduction

1.1.1 Outline

This treatise covers many topics such as radiation interactions, the physics of semiconductors, electronics and medicine. All of these items are linked by the desire to replace medical X-ray film with an improved system that will reduce radiation dose, improve diagnosis and possibly lower costs. In this chapter, the principle elements that together make a good medical X-ray detector are discussed and a design for an improved system is presented. Chapters 2, 3 and 4 concern themselves with the operation of the proposed a-Si sensor pixel. Chapter 5 concludes with a discussion of the results obtained.

1.1.2 Amorphous Silicon and Large Area Electronics

Despite its many disadvantages, compared to crystalline silicon (c-Si), amorphous silicon (a-Si) has found many uses as a semiconductor. Two principle properties have driven the recent interest in this material.

Firstly, is the ability to coat large areas with thin films of a-Si. This allows the production of items like the digital active matrix flat panel displays, which make up a large proportion of today’s laptop PC displays. Such films can be grown with a high uniformity. Though clearly not as uniform as epitaxy, which has nearly atomic flatness, it is sufficiently accurate for most production needs, having variations in
thickness of a few percent across a substrate. For c-Si, the largest wafer dimension available at present is 12" at a cost of ~$1,000. The more common 8" wafer costs much less (~$110 per wafer) but neither of these really provide the area required for applications such as displays or a medical X-ray sensor. However, a-Si active matrix displays with a diagonal screen size of 12.1 inches are commercially available and single panel a-Si displays with a diagonal dimension of approximately 28 inches have been fabricated\(^1\).

There appears to be no difficulty in depositing even larger areas of a-Si, the real limit being on component yield. A 12" active matrix display will have approximately 1 million transistors\(^2\), and if one fails then the display is usually rejected. For this reason a-Si technology tends to be as simple as possible to help improve yield.

The second point of importance, which is common to all amorphous semiconductors, is that in theory they all act like direct band gap semiconductors. A direct band gap semiconductor allows the emission and absorption of energy from a photon without the need for a phonon interaction to carry away crystalline momentum (k). Crystalline silicon however, is an indirect band gap semiconductor and requires a phonon interaction (i.e. a change in k) to absorb or emit any light. For this reason, c-Si is of little use in opto-electronic applications. Hence, the increasing popularity of GaAs, which is a direct band gap semiconductor, and is thus better suited for optical devices.

In amorphous materials, k is no longer a conserved quantity\(^3\) as the periodic lattice which characterises a crystal is not present. Consequently these materials interact with photons easily and are good absorbers. This property makes a-Si ideal for solar cells. A thin layer of a-Si can absorb light far better than a thicker layer of c-Si across
the whole range of the visible spectrum. Combined with the ability to coat large areas cheaply this makes a-Si one of the principle materials for solar cell manufacture today. Mixing these two properties one can easily see that a-Si is an ideal material with which to construct large area arrays of photodetectors. Though it is an obvious choice its electronic properties are inferior to c-Si, but as will be discussed in Chapter 2, this does not present a problem for our application. In addition, the CCD is found in many applications because it is a readily manufactured component. Though physically small compared to an a-Si array systems utilising CCDs can be designed to use optics to reduce the image, although there is a consequent loss of image quality.

However, there are applications for which the need for optics is either impractical or an unnecessary expense. For example, a portable photocopier or fax can be made if direct contact copiers can be made\(^4\), but this requires a large area sensor. It is in such specialist markets that a-Si:H large area photodiode arrays may be able to make an impact.

The research contained in this thesis is primarily concerned with optimising a-Si:H photodiodes for use in medical image sensors, in particular as a replacement for X-ray film. Optimisation of photodiodes involves improving one aspect of the detector often at the cost of another. For example, to increase the rise time of any photocurrent generated in a photodiode one may choose to make the detector thinner, hence reducing the distance that charge (i.e. signal) has to travel to escape from the detector. However, the thinner detector now has reduced detection efficiency, as there is less material to stop the light. Consequently, before optimisation can begin, it is necessary
to target the application (i.e. medical X-ray imaging), decide what is required and then see how much can be achieved and what can be improved on.

Section 1.2 - Medical X-ray Imaging

In 1886 Roentgen discovered X-rays and found an immediate application for them by photographing his wife's hand. Since then, x-rays have been consistently used to allow doctors to see into the body. Though technology has led to refinements, the basic equipment is the same as that used by Roentgen over 100 years ago. However, since Roentgen's time we have come to understand that X-rays can have harmful effects, particularly in large doses. Currently it is estimated that 14 % of the UK population's lifetime dose of radiation comes from medical X-rays. This figure does not include the large radiation dose given to some patients as part of intensive radiotherapy but it is still quite high. Figure 1-1 shows how this dose compares to the dose from other sources of radiation. It is clear that medical exposures dominate the man made contribution to the radiation dose.

Figure 1-1: Breakdown of the Source of Radiation Dose to the UK Population.
Clearly any means by which the dose from X-rays is reduced can only be beneficial for the individual patient. The most important factor in achieving this is to have a detector with high collection efficiency. That is as much of the signal as is possible (X-rays in this case) must be collected (i.e. stopped in the detector) and the resulting signal extracted. X-ray films will typically stop 2% of incident X-rays directly and with fluorescent screens this efficiency can be increased to 50%; our system should exceed this.

Replacing the photographic plates in X-ray equipment with something that has a linear response would also help to achieve this. To see why this is consider Figure 1-2 which shows how the optical density of a film (which is proportional to the number of silver grains that precipitate during development) varies with X-ray exposure. Such a characteristic curve (also known as a H-D curve), often can be fitted by an equation of the form given in equation (1-1).

\[ D = D_{\text{max}} \left(1 - \exp(-kX)\right) \]  

(1-1)

where \( D \) is the optical density (defined at the log of the ratios of the incident to transmitted light), \( D_{\text{max}} \) is the maximum optical density possible for the film, \( k \) is an arbitrary constant and \( X \) is the exposure in arbitrary units.
Figure 1-2: A typical characteristic curve for X-ray film.

It can be seen that at low intensities there is no response to incident radiation. Then over a narrow range of intensities there is a linear response prior to the saturation of the film. X-ray images whose exposure is the region of linear response will be good quality images. If a radiologist fails to estimate the dose arriving at the film correctly then the film will be under or over exposed and the study will need to be repeated. It is useful in radiology to define a quantity - $\Gamma$, which describes the slope of this linear region - $\Gamma$ is defined in equation (1-2) below. The smaller the value of $\Gamma$, the larger the working range of exposures that can be obtained. However, the consequence is a lower contrast between exposures. Typically $\Gamma$ lies between 2 and 3. It is conventional to write $\Gamma$ as:

$$\Gamma = \frac{D_2 - D_1}{\log_{10}(X_2) - \log_{10}(X_1)}$$  \hspace{1cm} (1-2)

where $D_2$ and $D_1$ are the optical densities at each end of the linear region and $X_1$ and $X_2$ are the exposures at each end of the linear region.
Note that what is important is how intense the X-rays reaching the plate are and this depends not only on the intensity of the X-ray tube but also on what is being imaged. A "thick" patient will require a higher minimum dose than a "thin" patient will. Unfortunately, the only way to judge the required intensity of X-rays is to use experience and to carry out an exposure.

Inevitably, the result is that some X-rays need to be re-taken because the image showed no signal. Thus, the patient receives an excess of dose. A sensor with a linear response overcomes this as a low or an excessive dose will still result in a recognisable image.

Following on from this, another possible way of reducing dose is to image process any images. The latter may allow an X-ray system to be adjusted during acquisition of an image. Hence, if an image is clear then the X-ray tube can be turned off possibly saving the patient extra dose. Image processing may allow contrast enhancement and other features that could save a patient having to undergo further examinations if a feature is unclear. Both of these possibilities come down to the need to digitise the image to allow high-speed data transfer for real-time* imaging and image processing. This raises the possibility of having computers assessing image quality and operating the X-ray tube at different voltages and for different times to reduce dose.

---

* Real-time is a subjective term. Here we mean a stream of images can be viewed on a monitor and appear to be moving smoothly. This means having frame rates of about 25 Hz, that of television.
Given that we wish to have a digital output from the detector the X-ray image must be collected and converted into an electronic signal. This means using some sort of semiconductor converter. However, semiconductors such as Si and GaAs have a very low stopping power for medical X-rays, which typically have energies in excess of 80 keV. A simulation of an X-ray spectrum as might be obtained from a medical X-ray tube is shown in Figure 1-3.

![X-ray Spectrum Simulation](image)

**Figure 1-3:** A simulated X-ray spectrum for a Tungsten anode X-ray tube with a 2.5 mm Al filter, an anode voltage of 100KV, an anode angle of 17° and an anode to sample distance of 1 m.

Over 21 mm of Si is required to have a 66% probability of completely stopping an 80 keV X-ray. To illustrate this, Figure 1-5 (a) shows the mass attenuation coefficient of Si over the energy range 0-100 keV, and Figure 1-5 (b) shows what fraction of the incident signal would be collected for detectors of varying thickness.

Typically, hydrogenated amorphous silicon (a-Si:H) is grown to a thickness of 1 micron, and is insufficient to stop enough X-rays for direct conversion. Growing thicker layers tends to lead to the a-Si:H peeling off the substrate because of the strain that occurs at the substrate/a-Si:H interface. This can be offset somewhat by using He
in the growth process. For example Pochet et al.\textsuperscript{6} use this process to build an ionising particle detector.

However, this is still an unproven technology and there are other possible ways of approaching this problem. The first way is to use a semiconductor with a higher nuclear charge (Z), and hence a larger linear attenuation coefficient in the X-ray region, than Silicon, for example CdTe. However, the growth of high Z semiconductors is still in its infancy and consequently this technology is as unproven as thick layers of a-Si, though it is promising and others are currently researching it.

A second approach is to convert the X-rays into lower energy photons by using a scintillator and then collect the photons using photodiodes. This thesis is concerned with this approach. In particular, we would wish to use CsI as the scintillator and a-Si:H as the photodiode converter. Figure 1-6 shows the mass attenuation coefficient of CsI and the absorbed X-ray flux, as a percentage. It is clear that CsI is far better at stopping X-rays than Si, and about 100 \textmu m of CsI is capable of stopping a sizeable fraction of X-rays (e.g. 50% of X-rays at 50 keV). Figure 1-4 illustrates this advantage by comparing CsI with Si.

There are other issues regarding the choice of scintillator. CsI has a light output that is well matched to the optical absorption of a-Si. Many scintillators peak in the region of blue to UV light and are hence difficult to detect with photodiodes. In addition, CsI is durable compared to some of its counterparts as it is soft and malleable. Finally, the possibility of growing CsI with columnar structures is of considerable importance in keeping the spatial resolution of the detector as high as possible. Inevitably when
light is generated, it will spread out equally in all directions, and a columnar structure can help "pipe" light towards the sensor.

![Graph comparing energy and X-ray absorption](image)

Figure 1-4: A comparison of values from Figure 1-5 and Figure 1-6 illustrating the improved absorption of X-rays in CsI relative to Si.
Figure 1-5: (a) The mass attenuation coefficient for Silicon and (b) the percentage absorption X-rays in Silicon of varying thickness, as calculated from the figures in (a) with the spectrum from Figure 1-3.
Figure 1-6: (a) The mass attenuation coefficient for CsI and (b) the percentage absorption of 80 keV X-rays in CsI of varying thickness, as calculated from the figures in (a) with the spectrum from Figure 1-3.
Section 1.3 - Detector Design and Principles

Our X-ray imaging detector will be based on the pixel design shown in Figure 1-7, and is called the integrating pixel detector. Here X-rays are incident on a thick (~300 µm) layer of columnar CsI(Tl), a scintillator which emits light with an average wavelength of 550 nm. The top reflecting contact and the columnar light pipe structure of the scintillator helps to reduce light spreading and enhance spatial resolution.

Figure 1-7: Schematic of a single pixel in the Integrating Photodiode Image Sensor

The photodiode at each pixel collects the incident light and converts it into a photocurrent. A storage capacitor beneath each photodiode then integrates this current. When the capacitance is much greater than that of the photodiode, an improved linearity in the response is achieved.

During readout of the imaging signal, the a-Si:H Thin Film Transistor (TFT) can be turned on by applying a suitable, positive, gate voltage. This allows the stored charge
to be passed to a conventional c-Si integrating amplifier and an Analogue to Digital Converter (ADC). The detail of reading out the signal from the a-Si photodiode will not be discussed here. Figure 1-8 shows how the TFTs for holding and releasing the signal are arranged. This system is simply based on the active matrix display used in many systems today.

![Diagram of a typical active matrix structure](image)

**Figure 1-8:** Schematic of a typical active matrix structure, pulses along the gate lines turn on the transistors causing the charge from the pixels to be read out along the data lines.

The problems of converting the analogue signal to a digital one, switching the TFTs and reading out the data has been studied by other workers. These systems are destined to be implemented as custom built ASICs which will keep electronic noise to a minimum. This is worth noting for when we return to the a-Si:H photodiodes later,
Elements of Design for a Digital Medical X-ray Imager

as it is important that any noise contribution that they might make is also kept to a minimum, such that the electronic noise of the readout dominates.

Whilst discussing the external electronics it is worth noting that on a large a-Si:H array it is inevitable that the electrical characteristics of both photodiodes and TFTs will be subject to some variation (due principally to process variation). It is expected that as part of the read out systems, a Digital Signal Processor (DSP) will be used to allow “gain correction”. This simply means that the sensor will be calibrated against a standard source. This will supply a DSP with calibration data to allow it to modify the signal from each pixel so that all of the pixels are corrected to show a uniform signal response.

Section 1.4 - Advantages and Disadvantages of the Integrating Pixel Detector

The basic advantage of this detector is the linearity of response we expect to achieve compared to the non-linear response of film screen systems. In addition, the digital output allows for rapid real-time imaging. This cuts out the need for time consuming developing of plates. Another beneficial consequence is that since the detector is reusable, unlike film, there is a reduction in consumables cost. (However, this saving is debatable, as there is the additional cost of disc and memory storage space). All of this leads to a more convenient system that will help reduce dose.

Because the system is specifically targeted as a replacement for X-ray film, it has no need for energy resolution. This is because of the integrating nature of the detector. A detector with energy resolution would have to be event driven, allowing a count of electrons every time an X-ray interacted with the system. Attempts to fabricate such
systems have been made\(^7\), but these rely on crystalline semiconductors and so suffer from small area and high cost. The reason for using crystalline semiconductors is simple; amorphous semiconductors are too slow. This would mean that having registered an event, a pixel would be dead for microseconds, long enough to lose a considerable amount of information. In addition, there is the question of reliability; the more complex the detector electronics then the greater the chance of failed pixels. This is not a problem for our purpose but it does emphasise how the application must be kept in mind at all times. Having said this, one advantage of the system is that by changing the scintillator to Gd, a neutron detector can be made and by replacing the photodiode with a gas chamber, a beta or alpha particle sensor can be made. Such systems, with their attendant difficulties, will not however be the principle concern of this thesis.

**Section 1.5 - Requirements for Photodiode Detectors**

Having fully described our system, we can now see what is required of the photodiodes. First, we want a high collection of light. This is best measured by the quantum efficiency (\(\eta\)) of the diode which is defined differently by various authors. For this reason, defining it will be postponed until chapter 2 where it will be discussed in more detail. However, simply put, the quantum efficiency is the number of electron-hole pairs generated when an optical photon is absorbed.

Note however, that not all of the charge generated will be collected; some of it will be trapped in the material. Because of this, it is worth making a distinction between the quantum efficiency, which measures how efficient the material is at converting
photons to charge, and the charge collection efficiency ($\chi$). The charge collection efficiency is simply the fraction of generated charge carriers that reach the contacts of the material, and so are detected. An ideal value for $\chi$ is 1, indicating that complete charge collection has been attained. A value greater than 1 would indicate some sort of avalanching process is occurring resulting in a photocurrent gain in the system.

Ideally, a photodiode would have a quantum efficiency of 1, but they rarely do because of the trapping of carriers by defects. In addition, quantum efficiency will vary with the wavelength of the incident radiation, where quantum efficiency is critically dependent on the band gap of the material. For the X-ray detector, we require a diode that has high quantum efficiency at the green wavelengths at which the CsI(Tl) scintillates. As well as a high quantum and charge collection efficiency, it is important to stop as many light photons as possible. The number of photons absorbed at a particular wavelength is given in equation (1-3) below.

$$I = I_0 \exp(-\alpha x)$$ (1-3)

where $I_0$ is the incident intensity of light, $I$ is the intensity at a depth $x$ (cm) and $\alpha$ is the optical absorption coefficient (cm$^{-1}$).

Figure 1-9 shows a typical plot of Absorption Coefficient, $\alpha$, vs. wavelength and from this, it can be found that a thickness of 1 $\mu$m of a-Si:H will stop about 99.9% of 620 nm (red) light. The stopping efficiency can be boosted simply by increasing the thickness of the a-Si, but as will be seen in chapter 2, this reduces charge collection efficiency.
Figure 1-9: A plot of Optical Absorption ($\alpha$) vs. Wavelength (from Street$^8$, pg. 89, After Cody$^9$ et al.)

Figure 1-10 shows how our detector would be integrated. It can be seen that the presence of the TFT, the data lines and the gate lines would make a significant fraction of the photodiode insensitive to optical radiation. Additionally, any charge generated by optical photons at the edge of the pixel may drift into the adjacent pixel. Consequently, pixels must be separated by a small gap or insulator. Both of these factors reduce the collection efficiency of an individual pixel and over an array of 10 cm by 10 cm; with an 80% fill factor, as much as 20 cm$^2$ of the array would be dead.
The design shown in Figure 1-11 has been proposed by Morton and colleagues to address these problems. By inducing a lateral field across the detector, it should be possible to set up a non-uniform drift field to focus current generated at the periphery of the diode to the central electrode. By placing such drift elements over the optically inactive elements of the underlying active matrix, we aim to solve the fill factor problem. This is because this focused photodiode can be built over the TFT. In addition, the possibility for charge generated at the edges of the detector to fall into adjacent pixels is reduced because the lateral field imposes a drift towards the centre.
of the detector, where it still can be collected. An analysis of the a-Si:H photodiode is presented in Chapter 3 where a wide range of options will be considered.

**Figure 1-11: A focussed photodiode structure proposed for reducing the fill factor in a-Si photodiodes.**

Reduction of cross talk is also an important issue. Whilst the detector's active area should be maximised, this cannot be done at the cost of allowing large signals to drift to adjacent pixels. It is important to realise that cross-talk will arise at many points in the detector. For example, there will be cross talk in the scintillator as light is transported onto other pixels and there could be cross talk in the readout of signal between the overlapping gate and signal electrodes.
The response speed of the photodiode is also important. For a 10 cm by 10 cm array with a 100 μm pixel size, there are 1 million pixels. To achieve a reasonable “real-time” image, all of the pixels should be read out over 0.05 seconds (i.e. 25 frames/second). At 1000 pixels per data row, this gives a readout time per pixel of 50 μsecs, assuming that all the pixels in a row can be read out in parallel. Clearly, the photodiode should have a response time better than 20 msecs to allow time for the signal to be placed on the storage capacitor between readout cycles. Of concern is the probability that signal charge will be trapped within the a-Si:H diode. The result is an image lag where information from the previous image is superimposed on the current image. Typical values of response time (a signal falls to < 2% in 40 msecs\(^1\)) indicate that this could be a problem in high rate applications. However, this is close to acceptable values and it would be hoped that these values could be lowered.

Finally, but possibly of most importance to this thesis, is the leakage current and the signal to noise ratio of the photodiode. We require that the signal from the photodiode is sufficient to be distinguished from any noise in the system. This includes the noise that will be added by the TFT switch. The most significant source of noise will be the leakage current from the p-i-n diode, which should be kept to a minimum. However, it should be realised that a large leakage current is not automatically bad if the signal to noise ratio of the detector is large compared to that inherent in the radiation beam. Conversely, a low leakage current is of no use if the signal is only slightly larger and so will be lost during integration and readout.
Section 1.6 - Summary

We have discussed the need for, and outlined the design of a digital X-ray sensor designed to replace medical X-ray film. From this initial discussion, the following features are seen to be desirable.

I. A high efficiency, i.e. as much signal per unit incident X-ray photon as possible.

II. Good charge collection from the photodiode over a short time period.

III. A large light collection area (or fill factor) and good absorption of light.

IV. Low cross talk.

V. A sufficiently high response speed to avoid image lag.

VI. A good signal to noise ratio.

In the next few chapters, we will focus on the theory and simulation of an a-Si X-ray detector.

Section 1.7 - References


Elements of Design for a Digital Medical X-ray Imager


Chapter 2 The Physics of a-Si:H Photodiodes

Section 2.1 - Introduction

Photodetectors are commonplace semiconductor devices and many texts exist which discuss their operation and construction. However some of this literature does not apply to amorphous silicon devices and much of the specialised literature on a-Si:H devices concentrates on solar cell applications. This chapter briefly describes how our devices are constructed, discusses the basic physics needed to understand their operation and reviews some of the past work carried out on p-i-n* structures.

Section 2.2 - The Production of a-Si devices

There are many techniques for depositing amorphous materials such as evaporation, sputtering and chemical vapour deposition (CVD). Elliot, in chapter 2 of his book\(^1\) discusses the outline of these and other techniques. However, for growing device quality a-Si, only one technique is of use and that is PECVD (Plasma Enhanced CVD), sometimes known as glow discharge deposition. Figure 2-1 shows the basic set up of a PECVD reactor, the bullets highlight the various deposition parameters that affect the type and quality of material grown, along with typical values and ranges for a-Si:H deposition.

* Throughout this thesis, the convention where the first listed item is the first deposited is used. So p-i-n means a layer of boron doped a-Si is deposited first, followed by a layer of intrinsic a-Si and finally phosphorous doped a-Si.
Figure 2-1: A schematic of a simple, RF plasma deposition system with the various parameters shown.

The basis of the technique is to use a large capacitor to couple an RF field to a gas, thus creating a glow discharge by dissociation of the gas molecules. The possible reactions that can result can be very complex, and depend on the various parameters given in Figure 2-1 and the chemical properties of the gas. Three simple reactions, for silane, are given in along with the energy required to drive these reactions.

\[
\begin{align*}
\text{SiH}_4 & \rightarrow \text{SiH}_2 + \text{H}_2 & 2.2 \text{ eV} \quad \text{(2-1)} \\
\text{SiH}_4 & \rightarrow \text{SiH}_3 + \text{H}_2 & 4.0 \text{ eV} \quad \text{(2-2)} \\
\text{SiH}_4 & \rightarrow \text{Si} + 2\text{H}_2 & 4.2 \text{ eV} \quad \text{(2-3)}
\end{align*}
\]

Clearly, as the RF power level changes, more energy becomes available for various reactions and the rate of the reactions varies. Consequentially, so does the type of material deposited. At higher powers, the reactions can in fact promote the growth of more complex molecules. However, to complicate things, the efficiency with which the RF field can break bonds, depends on the impedance matching of the field to the
gas. Consequentially, changing the RF frequency can also alter the properties of the film. Legal restrictions on the use of RF fields do however mean that the RF frequency is usually limited to 13.6 MHz, as at other frequencies emissions can interfere with communication channels.

An RF field, however, cannot sustain a glow discharge by itself. It is necessary for electrons and gas molecules to be excited and distribute energy via collisions to maintain the discharge. This requires that the gas be kept at a low pressure otherwise the glow discharge will be extinguished. Reducing the pressure, however, does have consequences with respect to the reaction rate in the gas, ion bombardment of the material and the growth rate. Also, there is the gas flow rate, which determines the time that molecules are in the chamber, hence how likely they are to disassociate.

With this multitude of variables, it can be seen that the process is complex. Because of the complexity of the interactions, glow discharge interactions are poorly understood and are the subject of intense study. Consequentially, the subject matter is too broad to be discussed in any detail here, but the interested reader is referred to the excellent books on the subject by Smith\textsuperscript{2} and Street\textsuperscript{3}.

Finally, having maintained a glow discharge, atoms and molecules will then diffuse from the glow discharge onto the sample where they will be deposited. However the arriving molecules may not bond to the surface but might damage it instead, removing atoms, or they may diffuse about on the surface before reaching a thermal equilibrium and bonding to the surface. This is a complex subject, more about it can be found in the above references and references therein. It is important to recognise that the rate of arrival of the precursors and the temperature of the substrate governs the nature of these surface reactions. Hence, this plays an important role in the type and quality of
the deposited material. For example, the hydrogen content in a-Si:H is very important and if the substrate temperature is about 280 °C then hydrogen will be driven out of the sample, reducing the quality of the material.

The reason that PECVD is the only technique suitable for growing a-Si devices is twofold. Firstly, the technique introduces hydrogen into the amorphous silicon which helps remove strain from the material and improves its electronic properties by reducing defect states in the band gap (see next section and Chapter 4 for more details). Secondly, by adding boron or phosphorous the material can be doped. This is done by adding di-borane (B₂H₆) and phosphine (PH₃) to the silane. Doping is vital for electronic devices and was first reported in a-Si by Spear and LeComber⁴. This discovery was a surprise, as theory predicted that doping of a-Si was not possible. To this day there is debate as to the detail of the doping mechanism but that described by Street in chapter 5⁴ of his book is widely accepted and this will be dealt with in the next section.

In addition to being the only useful technique for producing a-Si devices, PECVD has other intrinsic advantages, which makes it useful. Of most importance to us is the ability to cover large areas with uniformly thick films. This is of importance to our work because without being able to coat large areas, an X-ray sensor would have to rely on optics which reduce the efficiency of the sensor. This is one of the driving forces behind a-Si development as can be seen by its applications such as solar cells and flat panel displays.

Like most CVD processes, PECVD tends to deposit conformal films, a useful feature when depositing on structures e.g. trenches. In such cases, the film will remain of a uniform thickness making it ideal for depositing insulating layers on FET's (the most
common example of this being the deposition of amorphous silicon nitride (a-SiNₓ) in TFTs). Finally, simply by changing the gas composition different materials can be produced. For example instead of using silane as the depositing gas methane (CH₄) can be used and the result is amorphous carbon (a-C) which can be engineered to be close in structure to diamond. This allows the doping of the material to be easily graded, the stoichiometry of films to be varied across the sample thickness and for different materials to be deposited on top of each other all in the same reactor. Such layers can produce interesting quantum effects and can be used to good effect in avalanche diodes.

The principle disadvantage of PECVD is the complex interaction of the various parameters that define the properties of the final film. Frequently, the interactions are poorly understood and it is common to get slightly different materials from different deposition reactors. Reactor contamination is another major problem. Many of the chemicals used are highly reactive and will stick to the chamber walls only to come off in the next deposition. This is a big problem with dopants, such as di-borane, which are only required in small quantities to create large differences in device performance. Thirdly, is that limited growth rates and the large amount of strain present in amorphous layers, aggravated by growing on substrates with different lattice parameters, limits the thickness of films that can be grown. Finally, silane is explosive at standard room temperature and pressure, and many of the other gases used are poisonous, flammable or explosive. This means that precautions must be taken to avoid serious accidents during deposition.

Having deposited a thin film on a suitable substrate, it can be then processed by the conventional semiconductor techniques, i.e. lithography and etching, to define device
structures before further layers are deposited. Ion beam implantation becomes unnecessary due to the ability to grow doped layers, carry out lithography and then grow further doped layers. However, annealing and deposition processes at high temperatures cannot be carried out on the samples. This is because at about 280 °C hydrogen will be driven from the sample. This degrades its properties and at higher temperatures, the structure will begin to crystallise.

Section 2.3 - From c to a: How amorphous materials differ from crystalline materials

2.3.1 Structure

From the content of many texts about solid state physics, one would be forgiven for thinking that all solids are crystals but this is far from the truth. A crystal is usually defined as a solid that has a basic unit, which can be repeated at periodic intervals to create a structure (i.e. it has translational symmetry). This basic unit (a primitive cell) may be a single atom or any arrangement of atoms that has translational symmetry. Such a system is useful, as the number of possible arrangements of the unit cell is restricted and can be deduced from symmetry arguments. This reduces an infinity of possible combinations down to a few possible structures. Furthermore, the repetitive nature of crystals means that they can now be analysed by powerful mathematical techniques such as group and Fourier theory. Of course, the ultimate power of defining a solid by a crystal is that by knowing the arrangement of the primitive cell the whole solid can be reconstructed, as can its electronic properties. It is therefore
very convenient that the lowest energy state of a solid is invariably a crystalline state; hence, most solids are crystals.

However, every time we look out of a window, we are reminded that not all materials are crystalline; window glass\(^*\) is probably the most common non-crystalline material. If you examine the arrangement of atoms in window glass you will find no repeating pattern; window glass is amorphous - that is to say without any repetitive structure on the atomic scale. It is important to recognise that when used in a scientific sense the phrase amorphous means a lack of structure on the order of the atomic scale otherwise the phrase amorphous solid is an oxymoron. After all a solid is defined by its ability to retain its shape. Compare this with a liquid, which will lose its shape without a container. Other common phrases for these amorphous materials are non-crystalline, disordered solids and glasses. It should be noted that glasses are a special type of amorphous material that undergo a distinct thermodynamic change when they solidify. All glasses are amorphous but amorphous materials are not always glasses. Because an amorphous material lacks any periodicity in its lattice, the material cannot be easily analysed theoretically. In contrast, the periodicity of crystalline materials allows an extensive theory of semiconductors to be developed. However, amorphous materials are not completely disordered (again this sounds contradictory but it depends on what scale you consider things to be ordered or disordered) and have some features in common with crystalline materials. In particular a-Si:H is

\(^*\) Note the use of the term window glass as opposed to glass. In a scientific context, glass has a much wider definition. Window glass is a silicon-oxygen based glass but it is possible to have glasses made from other materials.
tetrahedrally bonded just like crystalline silicon. However, bond lengths and angles will vary slightly throughout the material to create a random structure. Some bonds are even left dangling without another atom to bond to. This means that a-Si has a short-range order similar to crystalline silicon when only the nearest neighbours are considered.

2.3.2 Band Theory of Semiconductors and the Density of States (DOS)

Perhaps the most powerful and important theory relating to semiconductors is the band theory of semiconductors (see Kittel\textsuperscript{5} for a typical description of this theory). Here the periodic nature of the lattice gives rise to energies at which electrons cannot exist, the ubiquitous band gap. This allows the charge carriers to be thought of as moving in energy bands separated by these band gaps. The charge carriers in semiconductors then fall into one of two types - those in bands nearly full, called holes and those in bands nearly empty, called electrons - leading to the concept of conduction and valence bands. Carriers in these bands are free to move throughout the material and so the electronic properties of the material depend on the ability of these carriers to move and their number. Therefore, we arrive at the classic band structure as shown in Figure 2-2(a).

The band structure of semiconductors arises from consideration of perfect crystals, but these would have to be infinite and have no defects or impurities to be correct. In reality, there are some impurities and defects in the crystal structure. Such defects alter the band structure slightly. These can be represented using individual states in the band gap (as marked by the dotted lines on Figure 2-2(a)). These states would be
spatially localised\textsuperscript{*}. Of these impurities the most commonly studied are the dopants Boron and Phosphorous which fall into the lattice and donate or accept an electron thus affecting the electronic properties of the material. Figure 2-3(a) shows how the dopant is forced into the lattice doping the material.

Moving to an amorphous structure there is no periodicity and so it might initially be expected that there would be no band structure. However, in the case of defects within the crystal, the perfect periodicity is disrupted but the band structure remains. From this, we can conclude that the existence of a band gap in a disordered material might be possible. How disordered is a material before this no longer applies? This is a difficult question but we can conclude from our glass windows that amorphous materials can indeed have a band gap. If window glass didn’t have a band gap energy greater than the energy of visible light then it would be opaque and our lives would be much darker.

Should a band gap exist, then what effect does the disorder have on the material? From the case of c-Si, it might be postulated that defect states in the band gap would appear. This is indeed what happens and these states are localised. But what is

\textsuperscript{*} Spatial localisation is simply where an electron (or hole) occupies a quantum state that has a real position in space (to within about 10 Å) from which the electron cannot move (except by quantum mechanical tunnelling) unless excited. This contrasts with the usual states in the semiconductor bands, which are extended, i.e. an electron occupying that state might exist anywhere in the material and so can conduct electricity. The quantum mechanics of localisation are complex. Further material on localisation can be found in Mott and Davis.
different is that instead of there being a few scattered states identifiable with lattice defects there is a distribution of states within the band gap as shown in Figure 2-2(b), a schematic DOS for a-Si.

Figure 2-2: Typical Energy versus Density of States plots for (a) c-Si illustrating the band gap and the presence of doping levels in the gap and (b) a-Si:H showing how states protrude into the band gap giving rise to the larger mobility gap.
Note first the similar features between Figure 2-2(a) and Figure 2-2(b); both have a conduction band and a valence band. Both have Fermi energies about half way between the band edges. However, note that the band edges in the amorphous case are marked as being ‘Mobility’ edges. What meaning does the apparently arbitrary mobility edge have? States within the mobility edge are localised. This simply means that electrons within such states are located at a distinct point in space and will remain there. This is in contrast to the states outside the mobility gap, called extended states, which do not have an exact location. That is to say, they pass through the entire material and an electron in an extended state may be found anywhere within the material. So electrons within the mobility gap are immobile and those outside are mobile (This is only true at absolute zero - above this electrons in localised states can tunnel to nearby states).

The presence of these localised states creates many of the differences between c-Si and a-Si. At low temperatures (~4 K) in c-Si, all electrons remain in the valance band and so no conduction occurs. In a-Si, electrons will remain in localised states and can tunnel from one localised state to another giving some weak conduction.

For our purposes, the role of such mid-gap states is as traps. Electron-hole pairs created by a photon interaction may become captured by these states and so become trapped for a considerable time. The result will be a loss in signal. As shall be seen, such traps have an effect on the formation of a semiconductor junction.

2.3.3 Doping of a-Si devices

There is one other major difference between c-Si and a-Si that has not been covered in the above discussion, that is doping. As discussed in Section 2.2 - a-Si can be doped,
a fact shown by Spear and LeComber. How does doping in a-Si work? Consider Figure 2-3(a) which shows a phosphorous atom acting as a dopant, the periodicity of the silicon lattice makes this the lowest energy bonding configuration for phosphorous. This is principally because for phosphorous to have a different bonding configuration would involve reducing or increasing its co-ordination number (i.e. the number of bonds that it has) and this would in turn affect the bonding of the silicon. Since there is more silicon than phosphorous it is energetically favourable for phosphorous to change its bonding structure to suit that of the lattice.

In an amorphous semiconductor, the lattice no longer exists to put the phosphorous into this type of bonding and the 8-N rule of Mott takes precedence. This says that any atom in a random network will attempt to prevent lone electrons by forming bonds or creating paired electron states. So an atom with valency, \( N \leq 4 \) will have \( N \) bonds and an atom with valency, \( N > 4 \) will have 8-\( N \) bonds. It follows from this that any group V dopants placed in a-Si will simply make three bonds and therefore not have any lone electrons to add as dopants to the material.

Street, on pages 160-168 of his book, describes a model that can explain the doping effect in a-Si in terms of dangling bonds. This model or variants thereof, is the one most widely accepted - though by no means the only one. In this model, there is a compensation of dopants by dangling bonds elsewhere in the material. Therefore, a phosphorous state with 4 bonds can be created. Instead of having an excess electron, which would violate the 8-N rule, this electron is captured by a dangling silicon bond elsewhere in the material. This violates the charge neutrality of the dopant but maintains charge neutrality in the material; a similar mechanism can be used to explain p-type doping with boron. Figure 2-3(b) shows how this might work. Note
that the dangling bond that compensates does not have to be near to the dopant. Also, note that phosphorous with three bonds is much more likely to form but there is some probability of the above reaction giving some doping. It is clear however that this doping mechanism is far from efficient which matches the observed low doping efficiency for a-Si.

Figure 2-3: The doping mechanism in Si, (a) demonstrates how P doping adds an addition electron to the material and (b) shows how a-Si achieves the same effect by using a dangling bond.
Section 2.4 - When opposites collide: The Semiconductor junction

2.4.1 Introduction

The Semiconductor junction is fundamental to the production of the transistor and the diode. It is therefore not surprising that most works on semiconductor devices study the p-n junction in considerable detail. (See for example chapter 2 of Sze’s book on semiconductor devices). Unfortunately the existence of mid-gap states in a-Si complicates the system and prevents a truly analytical solution of the semiconductor equations that is possible for c-Si. It is however possible to develop these equations and solve them numerically. This we shall do over this chapter and the next.

First, we shall examine some of the basic physics of the p-n junction as it applies to c-Si. This shall then be advanced to the p-i-n diode. This has no analytical solution but is important with respect to a-Si where p-i-n diodes are far more prevalent. The reason for this is that for use in photodiodes a-Si, p-n junctions are very poor light converters.

An explanation of this is that light generates electron-hole pairs that must then be separated and removed from the material. In c-Si, generated carriers can diffuse over distances of several hundred microns and so can easily diffuse over the p-n junction

\[ L_p = \sqrt{D_p \tau_p} \]

and the electron mobilities of c-Si (1500 cm²/V-s) and a-Si (10 cm²/V-s) along with carrier lifetime of 1x10⁻⁷ s. This gives a value of ~120 µm for c-Si and ~10 µm for a-Si. The real figure for a-Si is likely to be far lower due to the complex nature of transport in these materials.

\[ \text{§ These values can be justified by using the equation for diffusion length } L_p = \sqrt{D_p \tau_p} \]

and the electron mobilities of c-Si (1500 cm²/V-s) and a-Si (10 cm²/V-s) along with carrier lifetime of 1x10⁻⁷ s. This gives a value of ~120 µm for c-Si and ~10 µm for a-Si. The real figure for a-Si is likely to be far lower due to the complex nature of transport in these materials.
without recombining and thus being lost. In a-Si, the carriers can only diffuse a few microns, probably far less when the multiple trapping levels are accounted for. Hence, carriers will recombine, not contributing a signal. With a p-i-n structure, the i-layer will have an electrical potential across it because of the space charge in the p and n regions. When carriers are generated, in the i-layer, they will be swept out of the device by the electric field and no longer rely on diffusion, hence improving the performance of the device.

2.4.2 The Basic Semiconductor Equations

To begin to study the properties of the p-n and p-i-n diodes it is necessary to use the basic semiconductor equations so that numerical and analytical solutions can be arrived at. We shall discuss the 1-D forms for ease of understanding and use, however in the next chapter some equations will be generalised to a 2-D form.

The first set of important equations is the continuity equations (see equations (2-4) and (2-5) below). These, when coupled with the current equations, (2-6) and (2-7), and Poisson's equation (see equation (2-8)) describe the current flow and electric field in the device.

\[
\frac{1}{q} \frac{dj_n}{dx} = G(x) - R(x) \quad (2-4)
\]

\[
\frac{1}{q} \frac{dj_p}{dx} = G(x) - R(x) \quad (2-5)
\]

\[
j_n = q\mu_n \left( n\xi + \frac{k_T}{q} \frac{dn}{dx} \right) \quad (2-6)
\]

\[
j_p = q\mu_p \left( p\xi + \frac{k_T}{q} \frac{dp}{dx} \right) \quad (2-7)
\]
\[
\frac{d\xi}{dx} = \frac{\rho(x)}{\epsilon}
\]  

(2-8)

where \(j_n\) and \(j_p\) are the hole and electron current densities respectively, \(x\) is the position, \(G(x)\) is the carrier generation rate, \(R(x)\) is the carrier recombination rate, \(n\) is the electron density, \(p\) is the hole density, \(\mu_n\) and \(\mu_p\) are the electron and hole mobility's respectively, \(\xi\) is the electric field, \(\rho(x)\) is the space charge density and \(\epsilon\) is the dielectric constant.

Also of importance in these calculations are the recombination and generation rates in (2-4) and (2-5). The simplest form of recombination is radiative or direct recombination where an electron in the conduction band recombines with a hole in the valence band. Compared to recombination through gap states this is small and is often neglected. A more complex form of recombination comes in the form of Auger recombination where an electron recombines directly from the conduction band to the valence band but emits an electron in the process instead of a photon. Again, this form of recombination is often neglected as its effects are small compared to recombination through gap states. Neither of these effects will be taken into account in the calculations in this thesis.

More common is recombination through the capture of carriers by defect states that can exist in the band gap (see Figure 2-6 for an illustration). These are often described by the Shockley-Read-Hall (SRH) equation. For a single trapping level, this equation can be expressed as:-

\[
R_{SRH} = \frac{(pn-n_i^2)}{\tau_p \left[ n + n_i \exp \left( \frac{E_F-E_i}{k_BT} \right) \right] + \tau_n \left[ p + n_i \exp \left( \frac{E_F-E_i}{k_BT} \right) \right]}
\]  

(2-9)
where $R_{SRH}$ is the recombination rate, $\tau_n$ is the lifetime of electrons in the state, $\tau_p$ is the lifetime of holes in the state and $n_i$ is the intrinsic carrier concentration for electrons, $E_t$ is the energy of the trapping level and $E_F$ is the fermi level. The lifetimes can be defined in terms of the capture cross-section ($\sigma$), the trap density ($N_t$) and the thermal carrier velocity ($v_{th}$) using equations (2-10) below. It is sometimes useful to consider the effective lifetime, $\tau_e$, which is taken as being equal to the denominator of equation (2-9).

$$
\tau_n = \frac{1}{\sigma_n v_{th} N_t} \quad \tau_p = \frac{1}{\sigma_p v_{th} N_t} 
$$

(2-10)

2.4.3 The Silicon p-n Diode and the Depletion Region

To understand the operation of silicon diodes first consider two separate blocks of p-type and n-type material. Figure 2-4 illustrates how the energy levels for these systems should look. When these regions are brought into contact current will flow until the Fermi levels equalise. Clearly, this requires a built in potential to exist between the p and n type silicon with the value given below.

$$
V_{bi} = \frac{1}{q} \left( E_s - q \left( V_n + V_p \right) \right) 
$$

(2-11)

Here $V_{bi}$ is the built in potential, $V_n$ is the energy between the Fermi level of the n type material and the conduction band and $V_p$ is the energy between the p type material and the valence band.
The Physics of a-Si:H Photodiodes

Figure 2-4: Band diagrams for n and p type silicon. $E_c$ and $E_v$ are the energies of the conduction and valence bands respectively. $E_f$ is the Fermi level where the superscript denotes the doping of the material (i is intrinsic, p is p-type and n is n-type). $V_n$, $V_p$ and $V_{bi}$ are as defined in the text.

To support such a potential difference there must be a region between the p and n type material depleted of carriers. It is assumed that the doping in both the n and p type material is uniform. Further, it is assumed that the carrier concentration in the depletion region is negligible. This is the point at which the first difference between a-Si and c-Si arises. In a-Si, there is a significant space charge density due to trapped charge. The charge density depends upon the recombination in the device. Hence, it couples the Poisson equation to the continuity equations and only numerical solutions become possible. Section 2.5.1 discusses this in more detail.

Co-ordinates are chosen so that the junction between the materials is at $x=0$ and the depletion regions range from $x=-x_p$ (in the p type region) to $x=x_n$ (in the n type region). These assumptions allow a simplified version of the Poisson equation to be written.

$$\frac{\partial^2 V}{\partial x^2} = \frac{\partial E}{\partial x} = \frac{q}{\varepsilon_s} N_D$$

for $0 \leq x < x_n$ (2-12a)
The Physics of a-Si:H Photodiodes

\[-\frac{\partial^2 V}{\partial x^2} = \frac{\partial E}{\partial x} = -\frac{q}{\varepsilon_s} N_A \]

for \(0 \leq x < -x_p\) (2-12b)

where \(E\) is the electric field, \(q\) is the electric charge, \(\varepsilon_s\) is the dielectric constant, \(N_A\) is the number of acceptors and \(N_D\) is the number of donors.

If it assumed that the electric field beyond the depletion region is zero then these equations can be integrated to give the electric field. Thus:

\[E(x) = E(0) + \frac{qN_D}{\varepsilon_s} x \quad \text{for} \quad 0 \leq x < x_n\]  

(2-13a)

\[E(x) = E(0) - \frac{qN_A}{\varepsilon_s} x \quad \text{for} \quad 0 \leq x < -x_p\]  

(2-13b)

where \(E(0)\) is the electric field at the interface. This can be calculated using the assumption of zero electric field when \(x=x_n\) or \(-x_p\) and is:

\[E(0) = -\frac{qN_D}{\varepsilon_s} x_n = -\frac{qN_A}{\varepsilon_s} x_p\]  

(2-14)

Further integration leads to the value of the built in potential. This can then be algebraically manipulated to give the formula for the width of the depletion region given in equation (2-15).

\[W = \sqrt{\frac{2\varepsilon_s}{q}\left(\frac{N_A + N_D}{N_A N_D}\right) V_{bi}}\]  

(2-15)

2.4.4 The Shockley Equation (The ideal diode law)

To derive the ideal diode equation it is first necessary to develop the concept of the quasi-fermi level. It is commonly assumed that the free carrier concentration in a semiconductor can be described by a Boltzmann distribution of the form shown in equations (2-16a) and (2-16b).
where \( n_0 \) and \( p_0 \) are the equilibrium electron and hole concentrations, \( n_i \) is the intrinsic carrier concentration, \( q \) is the modulus charge on an electron, \( E_i \) is the intrinsic Fermi level and \( E_f \) is the Fermi level.

When a voltage is applied to a semiconductor then these equations are no longer strictly accurate as the Boltzmann condition of equilibrium no longer applies. However, it is useful to keep these equations, so it is assumed that it is possible to replace the Fermi level with some function that allows the Boltzmann equations to apply even under these non-equilibrium conditions. This function is the quasi-fermi level and so the free carrier concentration in a system with an applied voltage is now given by equations (2-17a) and (2-17b).

\[
\begin{align*}
n &= n_i \exp \left[ -\frac{q(E_i - E_f)}{k_B T} \right] \\
p &= n_i \exp \left[ \frac{q(E_i - E_f)}{k_B T} \right]
\end{align*}
\]

where \( n \) and \( p \) are the electron and hole concentrations, \( n_i \) is the intrinsic carrier concentration and \( \phi_n \) and \( \phi_p \) are the quasi-fermi levels for electrons and holes respectively.

Separating the Fermi levels out in this way allows the Boltzmann equations to be applied to systems under an applied bias. It is usual then to work with the quasi-fermi levels that can be defined from equations (2-17a) and (2-17b) as:-

\[
\begin{align*}
n &= n_i \exp \left[ -\frac{q(\psi - \phi_n)}{k_B T} \right] \\
p &= n_i \exp \left[ \frac{q(\psi - \phi_p)}{k_B T} \right]
\end{align*}
\]
\[
\phi_n = \psi - \frac{k_BT}{q} \ln \left( \frac{n}{n_i} \right) \tag{2-18a}
\]
\[
\phi_p = \psi + \frac{k_BT}{q} \ln \left( \frac{p}{n_i} \right) \tag{2-18b}
\]

From equations (2-18) and the product of equations (2-17) it can be seen that the electron density at the depletion layer edge (in the p type material) is:

\[
n_p = \frac{n_p^2}{p_p} \exp \left( \frac{qV}{k_BT} \right) = n_{p0} \exp \left( \frac{qV}{k_BT} \right) \tag{2-19}
\]

where \(V = \phi_p - \phi_n\), \(n_{p0}\) is the equilibrium electron concentration in the n-type material and \(n_p\) is the non-equilibrium electron concentration.

Having established a method for calculating carrier concentrations, it is now worth considering the continuity equations (2-4) and (2-5). By substituting in equations (2-6) and (2-7), and noting the requirement that the carrier concentration remains in a steady state, equations (2-20a) and (2-20b) can be arrived at. If these equations can be solved to give the carrier concentration then the current-voltage relationship can be derived.

\[
G(x) - R(x) + \mu_n E(x) \frac{\partial n_n}{\partial x} + \mu_n n_n \frac{\partial E(x)}{\partial x} + \left( \frac{k_BT}{q} \right) \mu_n \frac{\partial^2 n_n}{\partial x^2} = 0 \tag{2-20a}
\]

\[
G(x) - R(x) + \mu_p E(x) \frac{\partial p_n}{\partial x} + \mu_p p_n \frac{\partial E(x)}{\partial x} + \left( \frac{k_BT}{q} \right) \mu_p \frac{\partial^2 p_n}{\partial x^2} = 0 \tag{2-20b}
\]

Here \(G(x)\) is the generation rate as a function of position, \(R(x)\) is the recombination rate function, \(n_n\) is the carrier concentration in the n type silicon and \(p_n\) is the hole concentration in the n type silicon. \(\mu_n\) is the electron mobility and \(\mu_p\) is the hole mobility.
Equations (2-20a) and (2-20b) can be manipulated algebraically to give a single equation, given below.

\[ D_a \frac{\partial^2 P_n}{\partial x^2} - \frac{n_n - P_n}{n_n/\mu_p + P_n/\mu_n} E(x) \frac{\partial P_n}{\partial x} + (G(x) - R(x)) = 0 \]  

(2-21)

\[ D_a = \frac{n_n + P_n}{n_n q/k_B T \mu_p + P_n q/k_B T \mu_n} \]  

(2-22)

Here the ambipolar diffusion coefficient, \( D_a \), has been introduced and is defined in equation (2-22).

For a-Si, the recombination and generation function is a complex term and, as has already been mentioned, couples to the space charge term in the Poisson equation. However, for c-Si, assuming a low injection of charge carriers, the generation-recombination rate can be approximated by equations (2-23). These can be demonstrated by applying the approximation of low injection (i.e. majority carriers dominate so that in n-type material \( n_{n0} >> p_{n0} \)) to equation (2-9).

\[ U_n = -\frac{P_n - P_{n0}}{\tau_p} \]  

(2-23)

\[ U_p = -\frac{n_p - n_{p0}}{\tau_n} \]

where \( U_n \) is the recombination rate in n-type material, \( U_p \) is the recombination rate in p-type material, \( \tau_n \) and \( \tau_p \) are the carrier lifetimes for electrons and holes and the other terms are as defined earlier.

Using the low injection assumption, and noting that in the neutral region there is no electric field, the continuity equation takes on the simple form shown in equation (2-24). Here the Einstein relation, \( D_n = (k_B T/q) \mu_n \), has been used.
\[ \frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{D_p \tau_p} = 0 \]  

(2-24)

This is a simple 2nd order differential equation and can be solved with appropriate boundary conditions. At the contact, the carrier concentration must be \( p_{n0} \), and at the edge of the depletion region, the carrier concentration is given by equation (2-19). Thus, solving for \( p_n \) and substituting into equation (2-7) gives \( J_p \) as:

\[ J_p = \frac{qD_p p_{n0}}{L_p} \left( \exp\left(\frac{qV}{k_BT}\right) - 1 \right) \]  

(2-25)

where \( L_p = \sqrt{D_p \tau_p} \) which is the mean free path length for holes.

Repeating the solution for the other junction and summing the results leads directly to the ideal diode equation, equation (2-26).

\[ j = j_s \left( \exp\left(\frac{qV}{k_BT}\right) - 1 \right) \]  

(2-26)

\[ j_s = \frac{qD_p p_{n0}}{L_p} + \frac{qD_n p_{n0}}{L_n} \]

2.4.5 Non-Ideal Diodes

The Shockley equation does not accurately predict the characteristics of photodiodes. This is because it neglects or simplifies several effects such as (1) Generation-Recombination, (2) High Current Injection, (3) Series resistance, (4) Band to Band Tunnelling and (5) cross talk or shunt resistance. Figure 2-5 illustrates an ideal diode curve versus a more realistic curve and the regions corresponding to the above effects.
Figure 2-5: A plot of a forward bias ideal diode characteristic compared to a more typical response illustrating the various components that cause deviations from the Shockley equation. The first of these effects arises because of the simplified form taken for the recombination/generation term in the diode. It is assumed, in the Shockley equation, that recombination is always dominant. When the diode is reverse biased this changes and as the carrier concentration in the depletion region drops, generation begins to dominate. This exhibits itself as an increase in the reverse bias current caused by carriers being generated in the device.

Equation (2-27) gives a formula for this current, here the SRH equation (equation 2-9) has been used along with the assumption that the carrier concentration is far less than the intrinsic carrier concentrations.

$$J_{gen} = \int_{0}^{W} q[G(x) - R(x)] dx \approx \frac{q n W}{\tau_e}$$

(2-27)

The reverse bias current is then the sum of the generation term (equation 2-27) and the diffusion term (equation 2-25). Assuming the voltage is greater than the thermal...
voltage \((k_B T/q)\) then the reverse bias current can be approximated by equation (2-28).

Note the presence of the depletion layer thickness in the generation term, the importance of this will become more apparent when discussing p-i-n diodes. In c-Si, the intrinsic carrier density is sufficiently low that the generation term can dominate. For a-Si, the effective carrier lifetime is sufficiently low to mean that this is also true. However, as is discussed later, the generation term in a-Si is likely to be far more complex.

\[
J_R = q \frac{D_p n_i^2}{\tau_p} \frac{q n W}{N_D} \frac{1}{\tau_e} \tag{2-28}
\]

In forward bias, as well as the diffusion current predicted by the Shockley equation, there is a recombination component. Using equation (2-19) in equation (2-9), the recombination current can be estimated to be:

\[
J_R = \int_{0}^{w} q Ud_x = \frac{q W}{2} \sigma v_{th} N_i n_i \exp \left( \frac{qV}{2k_B T} \right) \tag{2-29}
\]

This leads to the more general form for the forward biased diode equation, as shown in equation (2-30), where \(n\) has been introduced as the diode ideality factor or emission factor and typically lies between 1 and 2.

\[
J_F \approx q \frac{D_p n_i^2}{\tau_p} \frac{1}{N_D} \exp \left( \frac{qV}{k_B T} \right) + \frac{q W}{2} \sigma v_{th} N_i n_i \exp \left( \frac{qV}{2k_B T} \right) \approx \exp \left( \frac{qV}{nk_B T} \right) \tag{2-30}
\]

The high injection effect occurs because in solving equation (2-21) we have assumed that the majority carrier concentration is close to the equilibrium carrier concentration. With increasing forward bias, the increasing current means that the carrier density moves sufficiently far from equilibrium to invalidate the Shockley equation. Under
extremely high injection, the carrier densities on either side of the junction will become equal. Using \( n_p = p_p \) in equation (2-19) leads to:

\[
J_p \propto n_p \approx n_i \exp\left(\frac{qV}{2k_BT}\right)
\]  

(2-31)

As the current is driven higher, the diode begins to act like an ohmic resistor. This is because though the junction is intact, a large amount of current is lost traversing the quasi-neutral parts of the diode outside of the depletion region. This is equivalent to placing a resistor in series with the diode and is purely a function of the bulk resistance of the material from which the diode is constructed.

The fourth effect listed for a non-ideal diode is band to band tunnelling. This quantum mechanical effect occurs at extremely high bias or doping conditions. Under such conditions the gradient in the band edges becomes so high as to allow carriers to go directly from the valence band to the conduction band or vice-versa. This exhibits itself as an extra generation term, which is sometimes modelled as an exponential, field dependent term (see equation (2-32)). Its effects are to generate an increased current at high voltages. Such current can be catastrophic, leading to junction breakdown from joule heating. For the purposes of this thesis, this effect is largely ignored. Its presence should be noted, however, as it is the limiting factor on the operating range of any diode.

\[
G(x) = AE'(x)\exp\left(-\frac{B}{E}\right)
\]

(2-32)

where \( A, B \) and \( \gamma \) are constants whose values depend upon the nature of the model used.
The previous effects are mostly material dependant and arise from the physics of the materials involved in making the diodes. Shunt resistance is mostly a device effect, as is cross talk. Both arise from the ability of current to bypass the diode's junction. Cross talk occurs when two adjacent diodes on a circuit find a conducting path from anode to anode or cathode to cathode that bypasses the junction. Shunt resistance occurs because the diode finds a conducting path around the junction. This shows up in the reverse bias characteristic as an increased current with an ohmic character. If the shunt resistance is very low then the reverse bias saturation is dominated by the resistor instead of the diode and all rectifying characteristics are lost.

2.4.6 The p-i-n Diode

To solve the p-i-n junction, we assume that the carrier concentration in the p and n regions is a constant and so we consider only the intrinsic layer. This is assumed to go from x=0 (where the n junction lies) to x=W (the p junction). A general form of the diffusion equation is given by equation (2-33), this is based on equation (2-21) and applying the low injection assumption.

$$D_n \frac{\partial^2 P_n}{\partial x^2} - \mu_p E(x) \frac{\partial P_n}{\partial x} + \frac{P_n - P_n^0}{D_p \tau_p} = 0$$  \hspace{1cm} (2-33)

If it is assumed that the electric field is constant then a general solution to this equation can be found of the form given by equation (2-34). Solutions can be found if E has a specific dependence on x (e.g. E=k/x results in a hyperbolic solution of the Bessel equation), however this is not particularly useful. The approximation of a constant field seems a reasonable starting place. For more complex field
dependencies, the Poisson equation should really be solved in parallel with these equations.

\[ p_n = A \exp(-\alpha x) + B \exp(\alpha x) + p_{n0} \tag{2-34} \]

Noting that at \( x=0 \) \( p_n = p_{n0} \), and at \( x=W \) we note that equation (2-19) still holds. Using these boundary conditions gives the constants \( A \) and \( B \) and the exponential constant can be found by solving the auxiliary equation of equation (2-33). The resulting values are given below.

\[ A = -B = \frac{1}{(e^{-\alpha W} - e^{\alpha W})} \left[ \exp\left(\frac{-qV}{k_b T}\right) - 1 \right] p_{n0} \tag{2-35} \]

\[ \alpha = \frac{-\mu_p E \pm \sqrt{(\mu_p E)^2 - 4D_p \tau_p}}{2D_p} \tag{2-36} \]

Differentiating gives the hole component and symmetry gives an equivalent set of equations for electrons at the p interface, leading to equation (2-37).

\[ j = qD_p p_{n0} \frac{(1+\alpha_p^2)}{\alpha_p (e^{-\alpha_c W} + e^{\alpha_c W})} \left[ \exp\left(\frac{-qV}{k_b T}\right) - 1 \right] + \]

\[ qD_n n_{n0} \frac{(1+\alpha_n^2)}{\alpha_n (e^{-\alpha_p W} + e^{\alpha_p W})} \left[ \exp\left(\frac{-qV}{k_b T}\right) - 1 \right] \tag{2-37} \]

This is clearly very similar to the Shockley equation, however, the reverse bias current term now has a field** and thickness dependence. If the electric field is equal to zero then \( \alpha_c = (D_p \tau_p)^{1/2}/L_p \) the carrier diffusion length.

** Note that as the diffusion coefficient tends to small values the current depends increasingly on the term \( \mu E \tau \), the mobility-lifetime-product.
To obtain a simple description of a photodiode we simply consider the incoming light to generate a uniform carrier density ($\Delta n$) that is added to the recombination rate. Because this term is independent of $x$, it cancels out except in equation (2-35) where it modifies the carrier concentration directly. Hence, we can express the photo current as the sum of the dark current term, $j_{\text{dark}}$, given by equation (2-37) plus an additional term $j_{\text{photo}}$ to give the total current.

$$j = j_{\text{dark}} + qD_n\Delta n\frac{(1+\alpha_n^2)}{\alpha_n(e^{-\alpha_nW} + e^{\alpha_nW})} \left[ \exp\left(\frac{qV}{k_BT}\right) - 1 \right]$$  \hspace{1cm} (2-38)

**Section 2.5 - The a-Si Diode**

2.5.1 Modifications to the Semiconductor equations

Throughout Section 2.4 allusions have been made to the differences between the simple c-Si diode and the a-Si diode. Until now, the detail of those differences has been ignored. This is because the form of the semiconductor equations for a-Si does not lend itself to an analytical solution. The reasons for this are (1) the presence of a significant amount of trapped charge in the semiconductor and (2) the complex recombination through multiple trapping levels.

Trapped charge in the semiconductor means that when solving Poisson’s equation the space charge is of the form shown in equation (2-39) below. This is simply a sum of the free carrier density, the trapped charge and the dopant density. A consequence of equation (2-39) is that in the material away from the junction, equations (2-12) no longer holds as the trapped charge term remains and its form is unknown.

$$\rho(x) = q[p(x) - n(x)] + [p_i(x) - n_i(x)] + [N_D(x) - N_A(x)]$$  \hspace{1cm} (2-39)
where \( p_t \) and \( n_t \) are the density of trapped holes and electrons respectively, \( N_D \) is the density of ionised donors and \( N_A \) is the density of ionised acceptors. 

To solve the equation (2-39), it is necessary to know the trapped charge; the SRH model can be expanded to calculate this charge. Simmons and Taylor observed that the SRH calculation can be generalised to the case of many traps at different energies. From an analysis of the capture and emission processes, it is possible to arrive at the following distribution function.

\[
\frac{f(E_t)}{e^E_t} = \frac{nC_n + e_p}{nC_n + e_n + pC_p + e_p}
\]

where \( f(E_t) \) is the probability of occupancy of a trap/state at an energy \( E_t \), \( C_n \) and \( C_p \) are the capture rates of the trap for electrons and holes respectively (given by the product of thermal velocity and capture cross section), \( e_n \) is the emission rate for electrons and \( e_p \) is the emission rates for holes.

With this method, it becomes possible to express the carrier densities, and the trapped charge densities, as shown below in equations (2-41) to (2-44). Clearly, the value for ionised donors and acceptors depends upon the doping.

\[
n = n_i \exp \left( \frac{(E_{F_n} - E_c)}{k_BT} \right)
\]

\[
p = p_i \exp \left( \frac{(E_v - E_{F_p})}{k_BT} \right)
\]

\[
n(E_t) = g(E_t)f(E_t)
\]

\[
p(E_t) = g(E_t)[1 - f(E_t)]
\]

where \( n_i \) and \( p_i \) are the intrinsic density of carriers in the conduction band and valence band respectively, \( E_c \) is the energy of the conduction band, \( E_v \) is the energy of the
valence band, \( E_{Fn} \) and \( E_{Fp} \) are the quasi-Fermi levels for free electrons and holes and 
g(\( E \)) is the density of states (DOS) at an energy \( E \).

Note that the values for \( n_t \) and \( p_t \) are for a specific energy and hence require integration across the trap energies to give the total trapped charge for equation (2-39). This means that a form for the density of states must be found. In addition, because of the presence of quasi-fermi levels, the trapped charge is now coupled to the free carrier concentration. This means that the Poisson equation cannot be solved without first solving the continuity and/or current equations. To do this the recombination/generation terms must be described for a-Si, as must the density of states. These items are discussed in the next two sections.

2.5.2 Recombination and Generation in a-Si

To do this, functions for the generation and recombination of carriers must be found. As discussed earlier, we will concentrate on SRH recombination as the dominant process in diodes. With the assumption that SRH statistics describe the carrier concentrations, then the recombination rate via states in the mobility gap is given by equation (2-45). That is simply that the sum of the recombination through each of the individual levels gives the total recombination. The first form of equation (2-45) is the discrete form, which is used in this thesis as numerical approximation to the second (continuous) form. Note that in contrast to equation (2-9) the terms are either subscripted or made functions of trap energy \( (E_t) \) to account for the many levels.
It is important to note that for a-Si the continuous trap distribution means that the integral form of equation (2-45) should be used. However, for numerical modelling only a finite number of states can be considered and so the summation form is used instead. Again, the justification for this generalisation can be found in Simmons and Taylor.\(^8\)

Unfortunately, using the Shockley-Read-Hall statistics as given above, neglects a possibly important form of recombination, that is recombination through dangling bond states. Consider Figure 2-3(b) where the doping mechanism of a-Si is shown. Each dangling bond can have two electrons, one electron or no electrons in it. Each such state is labelled according to the charge state of the atom associated with it. Hence a bond with two electrons makes the atom negative and it is referred to as a D\(^{-}\) state, an atom with one electron in the dangling bond is neutral and is denoted by D\(^{0}\). Neutral states correspond to the states considered in the above treatment of SRH recombination. Finally, a D\(^{+}\) state is an empty bond with no electrons. Clearly a D\(^{-}\) state can only combine with holes and a D\(^{+}\) state can only combine with electrons, otherwise more than two electrons (or holes) could occupy the same state.

Figure 2-6 shows how SRH recombination is viewed in terms of a band diagram. The electron and the hole recombine by being captured in a trap. However, if the charge of the state is considered, the situation changes. Initially it may seem that charged
states simply have a different lifetime than that of a neutral trap. Since they are usually shallow, it might be expected that the capture and release rates are high for one carrier but very low for the other sort of carrier. Hence, it is highly unlikely that any recombination will occur between these states. Therefore, they can be neglected as recombination centres and their only effect would be to reduce the number of free carriers at any time. This could easily be simulated with SRH statistics by simply adding another set of terms to equation (2-45) to introduce a separate distribution with different trapping lifetimes. Unfortunately, this assumption does not appear to be valid and this simplification will be returned to in chapter 4.

What actually happens can be seen in Figure 2-6(b). The carrier is captured by a charge defect state, which then becomes a neutral state. This now can have a different cross-section and may have a different energy arising from the Coloumbic interaction between the carriers in the defect. The change in energy of the state caused by this interaction is known as the correlation energy. Now the carrier may have a higher probability of recombining and hence a lower probability of re-emission to the conduction band.

Consequentially the SRH statistics are no longer valid as they assume that carriers recombine through states at a single energy. A better model will consider the distribution of the charged and uncharged defect states. This has been done by various researchers, although a full description is quite complex. A simplified version is carried out by Hubin et al.9,10 who show that by assuming a large generation rate, capture cross sections independent of energy and a steady state, the recombination function is given by equation (2-46).
\[
R_{DB} = \left( \frac{p}{\tau_{p0}} + \frac{n}{\tau_{n0}} \right) \left( \frac{n\tau_{n^+}}{p\tau_{p0}} + 1 + \frac{n\tau_{p^-}}{p\tau_{p0}} \right)
\]  

(2-46)

where \( R_{DB} \) is the recombination rate through charged states, \( \tau_{p0} \) is the capture time of holes in an uncharged state, \( \tau_{n0} \) is the capture time of electrons in an uncharged state, \( \tau_{n^+} \) is the capture time of electrons in a positively charged state and \( \tau_{p^-} \) is the capture time of holes in a negatively charged state.

Figure 2-6: The possible recombination mechanisms in a-Si, (a) the traditional SRH recombination through a single neutral state and (b) recombination through an amphoteric charged state with a correlation energy.

Having covered recombination, the other item required for the solution of the continuity equations is the generation term. Thermal generation of carriers has already been dealt with indirectly by the use of SRH statistics or Boltzmann statistics. This however, is not the only source of carrier generation available. Of principal interest to us is the optical generation of carriers.
A useful figure of merit to define is the quantum efficiency, $\eta$. Unfortunately, there are several definitions of quantum efficiency according to the usage. To avoid confusion we define the intrinsic quantum efficiency, $\eta_i$, as the number of electron-hole pairs generated per absorbed photon. This definition is usually used in solid state physics. Detector physics uses a slightly different definition, the number of detected electron-hole pairs per unit incident photon. This we refer to as the extrinsic quantum efficiency, $\eta_e$. It can be related back to the intrinsic quantum efficiency by considering the charge collection as will be shown in equation (2-50).

The number of absorbed photons in a material can be found using Beer's law relating the absorbed light intensity to the position in the material. So the generation rate can be calculated using equation (2-47) given below.

$$G = \eta_i \frac{P \lambda}{hc} \alpha(\lambda) \exp(-\alpha(\lambda)x)$$

(2-47)

where $G$ is the generation rate, $P$ is the incident optical power, $\lambda$ is the wavelength of the incident light, $h$ is Planck's constant, $\alpha$ is the absorption coefficient of the material and $x$ is the cumulative path length of the light.

It is worth noting the work of Kramer and van Berkel\textsuperscript{11} who identify another source of carrier generation in a-Si, that is Poole-Frenkel emission. Poole-Frenkel emission is the field assisted tunnelling of carriers from defect states into the conduction band. This is a distinct possibility for $D^0$ traps but it is not possible in $D^+$ or $D^-$ traps. Hence, the process would also require a thermal tunnelling from the resulting charged trap to complete. How such a process might occur is illustrated in Figure 2-7. Because such a process would generate new carriers in proportion to the square root
of the applied voltage, it would be expected that this would introduce the same
dependence in the reverse bias leakage current.

This process has not been modelled because of the complexity of the terms involved.
However, it is worth mentioning that in their work, Kramer and van Berkel conclude
that to model this process, the effective mass of carrier must be lower than measured
by other means. This is based on using the dc dielectric function, as is shown by
Hartke\textsuperscript{12}, the dielectric function that should be used is that of the material at high
frequencies. This is because as the carrier moves away from the trapping centre the
free charge moves to shield the emitted carrier from the trapping centre. This process
occurs over the same time frame as optical frequencies. Determining an exact
frequency for the process is not however possible.

There are several other possibilities for generation mechanisms in a-Si, but none of
these are relevant to the operation of the photodiode. Band to band tunnelling and
avalanche generation may well be relevant to the breakdown of the device but will
play no role in its operation.
2.5.3 The a-Si Density of States

Because of the lack of a definition for the density of localised states, which is required for the calculation of $n_t$ and $p_t$, it is impossible to solve equation (2-45). Measuring the density of states of amorphous semiconductors is of prime importance and there is no single way to determine it unambiguously (see Street$^3$, section 4.2 for a review of some experimental techniques used). For example, Figure 2-8 shows the comparison of various density of states as measured using different techniques. However, common features do exist, allowing some models to be suggested.
Figure 2-8: The density of states as determined by various methods. These being SCLC (Space Charge Limited Current), FE (the Field Effect) and DLTS (Deep Level Transient Spectroscopy) (From Street\textsuperscript{1}, pg. 121, after MacKenzie et al.\textsuperscript{13})

Usually the band tails can be represented by an exponential function and the doping defects can be shown as a Gaussian peak as illustrated in Figure 2-9. However, a simpler, and frequently more useful, form used by Hack and Shur\textsuperscript{15} is given in equation (2-48) and illustrated in Figure 2-10. This neglects the gaussian band of states due to the dopants.

Figure 2-9: A commonly used structure for the density of states using exponential band tails with a Gaussian density of states for the doping effects.
As can be seen in Figure 2-9, there are only a few gaussian states compared to the band tail states. It is therefore reasoned that they will have little impact on recombination current and hence the operation of the device. This is a reasonable assumption for the neutral defect states. However, as discussed in the previous section, the dangling bond model illustrates that the neutral states combine with the charged states to give an increased recombination rate. Unfortunately, given the restrictions of our modelling software we shall use this simpler form in most of our calculations. In Chapter 4 we will consider the origin of the more complex density of states shown in Figure 2-9, and how it may be used in calculations will be discussed.

\[ g(E) = g_{D\text{\,min}}(E) \exp\left(\frac{(E - E_{\text{mv}})}{E_D}\right) + g_{A\text{\,min}}(E) \exp\left(\frac{-(E - E_{\text{mv}})}{E_A}\right) \]  

(2-48)

where \( g_{D\text{\,min}} \) is the minimum density of donor states, \( g_{A\text{\,min}} \) is the minimum density of acceptor states, \( E_D \) and \( E_A \) are characteristic energies for the donor and acceptor tail states respectively, \( E_{\text{mv}} \) is the energy of the minimum of \( g(E) \) with respect to the valance band and \( E \) is the energy with respect to the valance band.

\[ \text{This obviously does not apply when the trapped charge density becomes sufficiently high to introduce space charge effects. Nevertheless, in reverse bias this assumption seems acceptable.} \]
Figure 2-10: Mid-gap Density of States used by Hack and Shur\textsuperscript{15} for simulation of a-Si photodiodes. In accordance with the paper we have taken $g_{Ami} = g_{Dmi} = 1 \times 10^{16}$ cm\(^{-3}\), $E_{mi} = 0.65$, $E_A = 0.053$ and $E_p = 0.088$.

2.5.4 Summary

In this section, we have derived the basic sets of equations that describe the physics of a-Si p-i-n diodes. As can be seen, they are complex but a little study will reveal that given a few physical parameters (principally the mobility, capture cross sections and the density of states) a solution could be achieved using numerical methods. Appendix A gives some examples of how such equations could be solved in principle, however for our purposes this is irrelevant as we shall use an existing software package to solve these equations in the next chapter. It is unfortunate that code this will not allow us to use all of the equations above in their fullest sense, but no numerical solution will ever be perfect. As long as we note what is missing, we can at least estimate where our model will fail. For example, the Poole-Frenkel emission will not be included so we may expect our results for reverse bias dark current to be
under-estimated. In Chapter 4, however, we will return to the issue of recombination models and attempt to develop an improved model.

Section 2.6 - Previous work on a-Si Diodes

Though a direct analytical solution of the semiconductor equation is not possible for a-Si diodes, approximations can be made that allow analytical solutions to be attempted. Such work is rare but there are a few papers worthy of note. An early attempt at such a solution was due to Crandall\textsuperscript{14} whose approach was to assume a constant field across the i-layer and a single energy for trapping states. With these approximations, Crandall was able to show that for weak illumination the current in the diode is given by equation (2-49).

\[ J = eG(l_n + l_p) \left[ 1 - \exp\left( -\frac{L}{l_n + l_p} \right) \right] \]  

(2-49)

where \( G \) is the generation rate of carriers, \( L \) is the diode length and \( l_p \) and \( l_n \) are the drift length for holes and electrons respectively. The drift length is defined as being the product of the carrier mobility, lifetime and applied electric field.

This suggested that the collection efficiency of a photodiode (in particular a solar cell) is dependent on the longer of the two drift lengths. Later however, numerical simulations (Hack and Shur\textsuperscript{15}) showed that the shorter of the two drift lengths was the limiting factor.

The second analytical solution of note was that of Hubin et al.\textsuperscript{9,10}. They showed that if the recombination function in equation (2-42) is used, with Crandall’s approach, then the shorter of the two drift lengths determines the collection efficiency.
Numerical solutions provide a simple way of solving the semiconductor equations. Hack and Shur\textsuperscript{15} have calculated the properties of a-Si solar cells using models similar to those outlined above. Figure 2-11 compares the results of an electric field calculation by Hack and Shur with the results of a time of flight (TOF) experiment by Konenkamp et al.\textsuperscript{20}.

In the experimental result, the field peaks at the p/i interface, this is often attributed to defects at the interface. Interestingly, Hack and Shur have not included these in their calculations, but they still observe a slight difference between the E field at the p and n interfaces. This depends on which side the illumination comes from.

This issue of a positive charge at the p/i interface will be returned to later. Firstly though, it is worth noting other interesting items in a-Si diodes that simulations have struggled to replicate. One is the observation of ideality factors greater than 2 (see McMahon and Madan\textsuperscript{16}). The second is the lack of any thickness dependence in the saturation current when the diode is reverse biased\textsuperscript{17,18,19}. The final issue is the exact nature of the reverse bias saturation current.
Figure 2-11: The electric field profile with an a-Si p-i-n device (a) as calculated (After Hack and Shur\textsuperscript{15}) and (b) as measured using time of flight techniques (After Konenkamp et al.\textsuperscript{20}).
Chen and Lee\textsuperscript{17} and others\textsuperscript{21,22}, associate values of $n$ approaching 1 with currents due to drift/diffusion of carriers in the material and values approaching 2 with recombination of carriers in the bulk. However, values of $n>2$ have been observed. Sichanugrist et al.\textsuperscript{23} suggests that this effect could be due to interface charge at the p/i layer. In addition, it has been shown in two papers (Pochet et al.\textsuperscript{24} and Matins and Forunato\textsuperscript{25}) that a 'lateral' current can be to blame for this effect. Figure 2-12 shows how this mechanism works. A current from outside the area of the electrode is induced which can flow to the top contact. When there is a highly conductive doped layer, then current can flow along the layer and out of the electrode. Hence, at low voltages, excess current is generated (as demonstrated in both papers) which can be removed by etching away the top doped layer.

However, McMahon and Madan\textsuperscript{16} eliminate the possibility of a heterojunction effect at the p-i interface or of surface channels in the diode. They propose that this behaviour can be explained by having more that one set of localised states, each with a different capture rate. This is intriguing, particularly in the light of the dangling bond recombination model of Hubin et al.\textsuperscript{9,10}. Clearly, two recombination paths can have different effective capture rates. This could be equivalent to the model proposed by McMahon and Madan\textsuperscript{16}. Following this logic, the supra-linear behaviour is a consequence of high defect densities and coupling between different charged defects.

The second item of interest is the independence of the IV characteristics of diodes on the diode thickness. Chen and Lee\textsuperscript{17} show that the recombination currents should vary with thickness (due to the larger volume for recombination). Matsuura et al.\textsuperscript{18} however suggest that this lack of thickness dependence might be explained by a current limiting layer which from experimental evidence is less than 770 Å thick.
They attribute this to a layer of defect states at the p-i interface. Further evidence for this effect was presented by Mittiga et al.\textsuperscript{19}.

Figure 2-12: The introduction of a lateral current, $j_l$, is illustrated here, showing how current from regions not under the electrode might contribute by having a high lateral conductivity. (After Pochet et al.\textsuperscript{24})

At this point, it is worth examining the suggestion of the existence of a positively charged layer at the p/i interface. Several papers have suggested this as a solution for supra-linear ideality factor and for the thickness independence of IV characteristics. However, the possible origins of such a layer is rarely discussed, for example, Lampert et al.\textsuperscript{26} give the following passage.

"The existence of a depletion region in the undoped layer close to the p/i-interface is generally accepted. Although the positive space charge is necessary for such a depletion region, its origin has not yet been sufficiently discussed."

One obvious source of such space charge would be the introduction of defect states in processing. However, this is unlikely, as over a decade of research into solar cell
processing in all likelihood would have produced a process that does not generate significant defects at the p/i interface. For example, inverting the deposition sequence does not introduce a significant defect density at the n/i interface.

There is however another, more likely source of this charged layer, and to understand it, it is necessary to once more examine the recombination. Because the density of acceptor like trapping states is higher than that of donor trapping states, it follows that there is an asymmetry in the recombination near to the contacts. Hack and Shur\textsuperscript{15} point out that many papers model the p-i-n junction as an i-layer with boundary conditions on the quasi-fermi levels to simulate the presence of the p and n layers. This they suggest leads to an error in the recombination rate near the interface. The full calculation reveals that there is a peak in the space charge density at the interfaces. It seems credible that this in fact the origin of the space charge at the p/i interface.

This author feels that there is no obvious physical origin for a set of positively charged interface states at the p/i interface. The explanation of the anomalous features of diodes (no thickness dependence and n>2) may be found in the complexities of the dangling bond recombination model. Because of the asymmetry in the density of states, it seems likely that in a p-i-n diode, the occupation of the D\textsuperscript{+} states becomes sufficiently high enough to create a limiting region of space charge. This would give rise to the above effects.

To conclude this section, it is worth mentioning the issue of reverse bias currents. Firstly, the reverse bias dark current does not saturate as would be expected for an ideal diode, or indeed a diode dominated by SRH recombination. Instead, there is a voltage dependence which, as Hegedus et al.\textsuperscript{21} note, has been fitted by both
exponential forms and power law forms. The source of this reverse bias current has not been determined yet, though Kramer and van Berkel\textsuperscript{11} argue that it could be due to a combination of the Poole-Frenkel effect and thermally assisted tunnelling.

Martins and Fortunato\textsuperscript{25} show that the reverse bias dependence of voltage on current could be the product of "shunt conductances" generated by unexpected conduction paths created during processing. Ultimately though there is still no definitive explanation as to why the reverse bias characteristics of some diodes show some voltage dependence.

\textbf{Section 2.7 - The Electronics of Photodetectors}

The previous sections have concentrated on the solid state physics that describe an arbitrary amorphous silicon photodiode. From these simulations, it is possible to calculate the Current-Voltage characteristics and similar electrical parameters. It is desirable, however, to consider a simpler model for these devices that can be used in circuit simulations and to define some figures of merit. This helps in making comparisons between devices.

The principle figure of merit for a photodiode is the extrinsic quantum efficiency, $\eta_E$, defined earlier. Intrinsic quantum efficiency can be simply related to the extrinsic quantum efficiency using a modified form of equation (2-47) as given in equation (2-50). (Note that pre-exponent $\alpha$ arises because of the integration across the absorbing region of the semiconductor).

$$\eta_E = \chi \eta_i \frac{P \lambda}{hc} \alpha(\lambda) \exp(-\alpha(\lambda)x)$$  \textit{(2-50)}
where $\chi$ is the charge collection efficiency and the other terms have been defined previously. The charge collection efficiency can be a complicated function of electric field, temperature, etc. but is simply defined as the ratio of the charge carriers generated within the device to the charge that reaches the device electrodes. This is also another useful figure of merit.

Ideally, we want the extrinsic quantum efficiency to be 1 or greater to produce the maximum signal. To do this we need the charge collection efficiency, the intrinsic quantum efficiency and the absorption coefficient to all be high. A-Si does have good intrinsic quantum efficiency and it has a good absorption coefficient (as do nearly all amorphous semiconductors) since it is effectively a direct band gap semiconductor.

Figure 2-13 shows a comparison of the imaginary part of the dielectric constant of a-Si with that for c-Si. This is related to the absorption coefficient and it is clear that for much of the spectra a-Si is superior to c-Si. Unfortunately, a-Si does not necessarily have a good charge collection efficiency due to trapping in the material. For this reason, we want to see high fields within the device to boost this charge collection efficiency.

However, in addition to the photo-signal generated by incident light there is always a background dark current (reverse bias saturation current) in photodiodes. Hence, a better measure than $\eta_E$ of the device performance of a photodiode is the photodiode gain. This is the ratio of the photocurrent** to the primary photocurrent and is given in equation (2-51).

** There are important distinctions to be made between photocurrent and light current. The light current is the current that flows through the device in the presence of a light
The Physics of a-Si:H Photodiodes

\[
\text{gain} = \frac{I_p}{I_{ph}} = \frac{I_p}{q \left( \frac{\eta E P \lambda}{\chi h c} \right)} \quad (2-51)
\]

Here \( I_p \) is the photocurrent and \( I_{ph} \) is the primary photocurrent that is defined as the number of charge carriers generated by the incident light. A related figure of merit is the responsivity of the diode (R), which is the ratio of photocurrent to the incident optical power.

![Figure 2-13: Imaginary part of the dielectric constants for c-Si and a-Si (taken from Street\textsuperscript{3}, pg. 86, after Pierce and Spicer\textsuperscript{27})](image)

The dark current of a device however is not the only signal that will degrade the signal. The diode has an associated resistance and so will generate Johnson noise. In addition, the recombination and generation of carriers in the photodiode will generate shot noise. Hence, it is important to include the effects of such noise when designing source. The photocurrent is that part of the current which flows because of the light source. Therefore, the photocurrent is the light current minus the dark current.
a practical diode. Thus, it is necessary to determine the signal to noise ratio for the diode. For our purposes a single diode will not suffice; it must be combined with many other diodes on an active matrix. Hence, it is useful to define an equivalent circuit for the diode to allow analogue circuit simulation of the device. Figure 2-14 shows an equivalent circuit for a focused photodiode structure that can be used within a circuit simulation (e.g. SPICE) to analyse the noise characteristics. From this circuit is possible to express the signal to noise ratio (the ratio of incident optical power to the output electrical power) as given in equation (2-52). It can be seen from this that obviously the ratio of photocurrent to dark current is critical, however. Another point of interest is that $R_p$ also influences the noise. This term, due to Johnson noise, can be reduced by either removing parasitic resistances (such as the load resistance of the transistor) or by reducing the bandwidth of the device (e.g. by placing an integrating component in the pixel).

$$\frac{(S/N)_{power}}{4q(I_p + I_N) + 8k_{B}T B / R_p} = \frac{(g \eta g P/hv)^2}{4q(I_p + I_N) + 8k_{B}T B / R_p}$$  \hspace{1cm} (2-52)

where $P$ is the incident optical power, $I_p$ is the photocurrent, $I_N$ is the current form noise sources (i.e. dark current), $R_p$ is the total parasitic resistance of the diode and $B$ is the bandwidth of the circuit.
Figure 2-14: An equivalent circuit for a photodiode. The principle element is still a diode, however, we add a shunt capacitance, \( C_s \), which represents the capacitance of the p-i-n structure and a shunt resistance, \( R_{\text{shunt}} \). The shunt resistance represents leakage across the diode element, the series resistance, \( R_{\text{series}} \), represents the voltage drop that occurs across the non-depleted regions of the diode. Finally, \( I_{\text{ph}} \) represents the photocurrent generated by the incident radiation.

**Section 2.8 - Summary**

In this chapter we have expressed the semiconductor equations (see (2-4) to (2-8)) and discussed how they can be formulated for the special case of an a-Si p-i-n photodiode. We have discussed the empirical data on a-Si p-i-n diodes along with the results of some simulations and found that certain anomalies in the behaviour of such diodes have yet to be satisfactorily explained in full, for example:-

- The invariance of the current with diode thickness (possibly due to a layer of interface traps).
- The variation of diode quality factor at low forward biases (possibly due to electrode geometry, interface traps or failure to consider dangling bonds).
- The voltage dependence of the reverse bias dark current (possibly due to failure to consider the recombination/generation terms fully).
• A conflict between those who show the majority carrier affects performance and those who show that the minority carrier affects performance.

We have compared various methods used in the past and noted that no single simulation has covered all of the possible interactions, which could give rise to the above anomalies. The effects that we note that should be considered are:-

• The presence of a mid-gap distribution of states affecting the space charge distribution and free carrier concentrations.

• The possibility of recombination centres (i.e. traps) at interfaces.

• Recombination of carriers through dangling bonds which gives a different form for the recombination current.

• The generation of carriers from dangling bonds via the Poole-Frenkel effect.

• The geometry of the electrodes used in making the devices.

Finally, we have discussed some of the common electrical properties that we need to derive from any simulation work for comparison with data from real devices.

Section 2.9 - References


The Physics of a-Si:H Photodiodes


Chapter 3 Simulation of Photodiode Structures

Section 3.1 - Introduction

Chapter 1 presented six important points that must be implemented in the design of an efficient X-ray detector. Most of these points are ultimately related back to material properties, however, cross talk and fill factor are both heavily influenced by the design of the device components. To improve these qualities, it is convenient to simulate the device components in an attempt to optimise their performance. In this chapter we will focus on the design of the photodiode in an attempt to optimise the fill factor of our device whilst maintaining a low cross talk.

At the heart of this chapter is the simulation of a class of devices that we call focussed photodiodes. These are diodes whose structures have been adapted to prevent 'dead' space in the diodes where charge is not fully collected. Any such dead space reduces the efficiency of the detector. Worse still if there is a non-uniform response across the detector's surface, then the detector will have non-linearities in its response. This becomes important if there are features in the signal with a spatial frequency higher than that of the detector.

The basic design for such a structure is shown in Figure 3-1. The principle of operation is that a basic p-i-n structure allows the collection of light from the active region between the anode and cathode electrodes. However, the introduction of focus electrodes should allow the introduction of a lateral potential that should allow some charge collection from the regions outside of the anode and cathode electrodes.
This chapter concentrates on the efforts made to simulate these structures using the SILVACO VWF package of process and devices simulation tools. The operation of these tools is not discussed in any detail, they simply find numerical solutions of the semiconductor equations given in chapter 2 (reference to equations (2-5) to (2-9)), Appendix A outlines how these equations can be solved numerically. Since defect states are only modelled by Shockley-Read-Hall statistics as discussed in chapter 2, they will have some limitations. These inaccuracies, however, should not prevent information about the fill factor and the operation of these devices to be calculated.
It is worth noting at this point that the SILVACO tools generate results for a two-dimensional device. Consequentially, some of the results may not be applicable to three-dimensional devices, but in general if the devices have appropriate symmetry solutions in two dimensions, these can be extrapolated to three dimensions. As an example of this, the SILVACO simulator produces current-voltage (IV) characteristics with current units of A/µm - to convert these to the more appropriate Amperes, they should be multiplied by the missing electrode dimension. To prevent confusion about how the third dimension is being formed, all answers are given in units of A/µm unless otherwise specified.

Section 3.2 covers a single basic structure used to determine the effect of the various material properties on device operation and to verify the models used. The structures of the focused photodiodes are then examined and the first simulations cover how the IV characteristics of the devices are expected to vary and the photocurrents observed in the devices. Section 3.4 discusses the measurements of the fill factor and the position response of the devices; possible improvements are also introduced. Finally, Section 3.5 summarises the results and discusses the accuracy of the measurements.

Section 3.2 - Basic Diode Properties

3.2.1 Introduction

The SILVACO tools are based primarily on c-Si with options to add new models to simulate other materials. In this section, we will examine the effect that the various models that we wish to use would have on a simple device. To do this, we shall
examine the IV curves of diodes simulated using various models, each model will progress from a simple c-Si diode to the complex a-Si photodiode.

The principle figures of merit that we shall use are the reverse bias dark current, series resistance, shunt resistance, ideality factor and where appropriate the signal to noise ratio. These figures of merit are chosen as quantities that are simple to extract and reveal information about the workings and suitability of the diode. It should be noted that in some cases the concepts involved are not necessarily valid. As an example, the series resistance is taken as being the resistance calculated in the ohmic region of the forward bias curve. It is expected that the forward bias series resistance of a p-i-n diode should exhibit some current dependence (see Sze, pg. 120). This means that in fact the defined series resistance is not the resistance of the intrinsic layer but that measured directly from the diode curves.

To clarify this point, the definitions used to calculate these values are given here. The reverse bias dark current is taken as being the current flowing through the diode at -5 V. This is also the assumed operating voltage of the photodetector. This definition avoids the thorny problem of defining where saturation begins in a diode and when leakage, recombination currents and shunt resistances take over. This is particularly relevant in a-Si diodes, where the reverse bias currents are dominated by recombination effects.

Series resistance is calculated by making a linear fit to the forward bias region of the dark IV curve, typically in the voltage range 2-5 V. In a similar way, the shunt resistance is estimated using the reverse bias region of the IV curve. Ideality factor is
determined by fitting the exponential part of the IV curve, typically in the forward bias region 0.1 V to 1.0 V.

The signal to noise ratio is measured by taking the ratio of the photocurrent to the dark current. Here we are defining photocurrent as being the excess current generated by the incident light.

Figure 3-1 shows a simple diode that might be used in a photodetector array. The device has a 100 micron pitch and is 1 micron thick, a good thickness for stopping the maximum amount of light. For the purposes of the simulations in this section, the diode is assumed to follow the structure shown in Figure 3-2 with the parameters given in Table 3-1. Note that the doping profiles are assumed to take the form of a constant value followed by a gaussian tail. This tail is defined by the gaussian which has a width, given by Table 3-1, when the doping concentration is equal to 1x10^{10} cm^{-3}. In addition, the silicon is assumed to have a background doping of 1x10^{10} cm^{-3} phosphorous atoms. This helps in the solution of the semiconductor equations, and in comparison to the doping in the contacts, it is effectively intrinsic material.

The results of the simulations that have been run in this section are given in Table 3-2. The simulations are labelled by an id number (i.e. ID#) for reference with the text. In addition, the conclusions as to what defines a model for c-Si versus a model for a-Si are summarised in Table 3-3 at the end of this section.
Figure 3-2: Structure and parameters of the diode used for simulations in this section. Only half of the structure is illustrated, the full structure is mirrored about the left edge to produce the final diode. The parameters are explained in Table 3-1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Diode Thickness</td>
<td>1.0 μm</td>
</tr>
<tr>
<td>2</td>
<td>Diode Width*</td>
<td>50 μm</td>
</tr>
<tr>
<td>3</td>
<td>Cathode Doping Width (i.e. thickness of n type layer)</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>4</td>
<td>Cathode Doping Tail (see text)</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>5</td>
<td>Anode Doping Overlap</td>
<td>2.0 μm</td>
</tr>
<tr>
<td>6</td>
<td>Anode Doping Width (i.e. thickness of p type layer)</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>7</td>
<td>Anode Doping Tail (see text)</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>8</td>
<td>Anode Width*</td>
<td>30 μm</td>
</tr>
<tr>
<td>9</td>
<td>Anode Thickness</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>10</td>
<td>Cathode Thickness</td>
<td>0.1 μm</td>
</tr>
</tbody>
</table>

Table 3-1: Parameters used to define the p-i-n diode structure of Figure 3-2.

* Since the structure is actually only a half of the full diode these parameters are only half of their final value.
3.2.2 The effects of Contact Sizing on Diode Parameters

The first experiment involved varying the anode contact size to determine the role that this would play in the diode's design. The forward bias and reverse bias curves of a 100 micron diode with anode lengths of 100, 60 and 30 microns are shown in Figure 3-3. These are referred to as ID#1, ID#2 and ID#3 in Table 3-2. As discussed previously, this table summarises the figures of merit for all of the simulations in this section.

It should be noted that the small bumps in the reverse bias characteristics are likely to originate in numerical errors in the simulation. At the low carrier concentrations using in reverse bias calculations it is inevitable that some divergence from the true solution will occur simply because of the large difference between the maximum and the minimum carrier concentration in the device.

Following from the discussion of diode parameters in section 3-2, it can be seen that we are only modifying the shunt resistance and the series resistance of the diode. It can be clearly seen that the change in contact size does not directly impact on the diode's operation as the ideality factor is very close to 1 in the exponential part of the curve. However, the forward bias (series) resistance increases with reduced contact size. This is as would be expected as resistance scales with contact area. Since this is a one-dimensional simulation, the increase in resistance should be inversely proportional to an increase in contact length and this is so. (Note however that for this to be so, the effective contact size, including the doped regions must be accounted for, demonstrating the importance of lateral doping at the contact).
Figure 3-3: Variation of Forward and Reverse Bias IV Characteristics with contact sizing.
Table 3-2: Basic Diode Properties extracted from the various simulations. See text for a description of these properties and how they are extracted.

The shunt resistance does not follow such a relationship as might be expected, rather it is inversely proportional to the contact length. Two things should however be noted. Firstly as discussed earlier, these figures of merit are defined loosely. In the device with a 100 micron contact there is no alternative current path and so this shunt

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100 micron contacts</td>
<td>1.07</td>
<td>413</td>
<td>1.99x10⁻¹²</td>
<td>2.77x10⁻¹²</td>
</tr>
<tr>
<td>2</td>
<td>60 micron contacts</td>
<td>1.07</td>
<td>672</td>
<td>3.24x10⁻¹²</td>
<td>1.70x10⁻¹²</td>
</tr>
<tr>
<td>3</td>
<td>30 micron contacts</td>
<td>1.07</td>
<td>13 K</td>
<td>6.28x10⁻¹²</td>
<td>8.80x10⁻¹³</td>
</tr>
<tr>
<td>4</td>
<td>[Mobilities (cm²/Vs)]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\mu_n=1000) (\mu_p=500)</td>
<td>1.08</td>
<td>676</td>
<td>3.24x10⁻¹²</td>
<td>1.70x10⁻¹²</td>
</tr>
<tr>
<td>5</td>
<td>[Mobilities (cm²/Vs)]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\mu_n=10) (\mu_p=500)</td>
<td>1.06</td>
<td>1.9 K</td>
<td>9.51x10⁻¹²</td>
<td>5.78x10⁻¹³</td>
</tr>
<tr>
<td>6</td>
<td>[Mobilities (cm²/Vs)]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\mu_n=1000) (\mu_p=1)</td>
<td>1.09</td>
<td>1.0 K</td>
<td>4.86x10⁻¹²</td>
<td>1.14x10⁻¹²</td>
</tr>
<tr>
<td>7</td>
<td>[Mobilities (cm²/Vs)]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\mu_n=10) (\mu_p=1)</td>
<td>1.09</td>
<td>92.0 K</td>
<td>4.42x10⁻¹⁴</td>
<td>1.25x10⁻¹⁴</td>
</tr>
<tr>
<td>8</td>
<td>SRH Recombination (\tau_n=\tau_p=1x10^{-7})</td>
<td>1.09</td>
<td>676</td>
<td>3.08x10⁻¹²</td>
<td>2.37x10⁻¹²</td>
</tr>
<tr>
<td>9</td>
<td>Optical and Auger Recombination</td>
<td>1.09</td>
<td>676</td>
<td>2.99x10⁻¹²</td>
<td>2.37x10⁻¹²</td>
</tr>
<tr>
<td>10</td>
<td>SRH with (\mu_n=10) (\mu_p=1)</td>
<td>1.34</td>
<td>92.0 K</td>
<td>3.55x10⁻¹²</td>
<td>6.79x10⁻¹³</td>
</tr>
<tr>
<td>11</td>
<td>Continuous DOS (\mu_n=10) (\mu_p=1)</td>
<td>1.14</td>
<td>16.4 K</td>
<td>4.75x10⁻¹¹</td>
<td>2.04x10⁻¹¹</td>
</tr>
<tr>
<td>12</td>
<td>Continuous DOS (\mu_n=10) (\mu_p=1)</td>
<td>1.58</td>
<td>3.7 M</td>
<td>2.00x10⁻¹²</td>
<td>1.18x10⁻¹¹</td>
</tr>
<tr>
<td>13</td>
<td>Continuous DOS (\mu_n=1) (\mu_p=1)</td>
<td>1.14</td>
<td>169.7 K</td>
<td>4.98x10⁻¹¹</td>
<td>1.18x10⁻¹¹</td>
</tr>
<tr>
<td>14</td>
<td>Continuous DOS (\mu_n=10) (\mu_p=1)</td>
<td>1.14</td>
<td>16.5 M</td>
<td>2.47x10⁻¹²</td>
<td>2.09x10⁻¹²</td>
</tr>
<tr>
<td>15</td>
<td>Continuous DOS (\mu_n=10) (\mu_p=1)</td>
<td>1.45</td>
<td>87 G</td>
<td>7.52x10⁻¹⁶</td>
<td>6.72x10⁻¹⁷</td>
</tr>
</tbody>
</table>
Simulation of Photodiode Structures

resistance must be due to some other effect. Given the large numbers involved (and hence the small currents) it is entirely possible that the resistance is due to numerical errors, and is in fact infinite.

A second point worthy of note is that the area of the shunt resistance is increasing, but its effective length is also increasing. For two dimensions, the increase in area is approximately linear, but the increase in length is also linear*. The result is an inverse relationship, in fact one would expect the shunt resistance to drop with decreasing contact size due to the quadratic increase in the area of the shunt resistor.

Both of these items could be an explanation, but both highlight the limits of simulations. The first illustrates that small changes have to be critically analysed before being accepted, as they can easily be numerical errors. From the second point, it can be seen that two-dimensional simulations can often miss three-dimensional effects and hence be unphysical.

3.2.3 The Influence of Mobility on Diode Properties

The first major difference between a-Si and c-Si, is that a-Si has a much smaller mean free path between collisions for electrons and holes. This can be represented as a lower mobility. In this set of simulations, the mobility for electrons was reduced to 10 cm²/Vs and that of holes to 1 cm²/Vs. These values are in line with typical values for a-Si found in the literature¹.

¹ This fact can be seen by considering the shunt resistor as lying along a line from the cathode edge to the anode edge. It then lies along the hypotenuse of a right-angled triangle whose length varies with that contact size.
Table 3-2 shows the results for these mobility simulations. Of immediate interest is the first mobility simulation, which is identical to the 60 micron contact sizing simulation (ID#2). However, the results do differ slightly, this gives some idea of how much error exists in these values due to numerical errors in the simulation and extraction procedures. Noting this, it is clear that the ideality factor is again unchanged with mobility. Only the series and shunt resistances are seriously affected by the changes in the mobility. It is noteworthy that changing either mobility has a similar effect on the resistances, but when both mobilities are changed, then the increase in resistance is nearly two orders of magnitude greater. However, examining the electric field and carrier concentrations produced by the simulations shows no difference between simulations 4 to 7.

The reason for this is that in forward bias the carrier concentrations in the intrinsic layer are approximately equal, constant and very high; this is illustrated in Figure 3-4. Because there are no gradients in the carrier concentrations then equations (2-6) and (2-7) suffice to describe the current flow. Initially the current is made of hole and electron components, with the electron current being 3 times the hole current. This difference arising because of the difference in mobilities. Reducing the electron mobility simply means that the hole component now dominates (by a factor of 50) and similarly reducing the hole mobility means the majority carrier has been changed. When both mobilities are lowered, the electron current is now 10 times the hole current but the electron current is now 150 times smaller, hence the large increase in resistance of two orders of magnitude.
Figure 3-4: Plots of carrier concentration versus depth in a forward biased p-i-n diode. These results are taken from simulation ID#7 with 5V on the anode. The n layer begins at x=0 and goes to x~0.2.
3.2.4 The influence of recombination on diode parameters

The next three simulations introduce recombination models into the simulations. The first simulation (ID#8) introduces a single energy trap level, as modelled by equation (2-9), the single level SRH equation. The carrier lifetime is taken as being 0.1 µs, which is considered typical for c-

Si$^2$. Simulation ID#9 also adds models for Auger recombination and optical recombination, as can be seen this makes no difference and justifies the statement made earlier, in chapter 2, that these can be neglected. As is expected, the presence of such a recombination centre pushes up the reverse bias current. The ideality factor however remains unchanged, this is because the recombination lifetimes are sufficiently low as to have little effect on the forward bias characteristics of the device. This is to be expected from device quality c-

Si as otherwise transistors would suffer a large loss in gain.

Moving to the next simulation (ID#10) both the SRH recombination and the low mobility models are used (individually in simulations 7 and 8). The series resistance is as expected the same as for the low mobility simulation.

It is expected that the dark current would increase, due to recombination. However, a lower mobility should decrease the dark current. As expected, the final result lies somewhere between the two extremes.

Finally, note that the ideality factor in the forward bias simulation has in fact increased. This may seem odd when compared to the failure for SRH recombination to affect the ideality in simulation 8. The reason for this is that with the lower mobility the diffusion of carriers is lower. Hence, the term due to carrier recombination in equation (2-29) is increasingly important.
3.2.5 The a-Si Recombination Model

The final step in these simulations is to introduce a continuous density of states as discussed in Section 2.5.3. For these simulations the continuous density of states given by equations (2-48) are approximated by a discrete set of 20 levels with the recombination calculated using equation (2-45). The parameters used in these equations are taken from the literature\textsuperscript{2,3} and are given in Table 3-3.

The first of this series of simulations (ID#11) uses the default c-Si parameters along with a continuous density of trapping states. For the first time, it is found that the forward bias region from 0.1 V to 0.5 V does not automatically fit the ideal diode equation. Instead it is necessary to reduce the voltage range, this is because the recombination distorts the forward bias characteristics in various ways. The ideality factors given Table 3-2 were extracted by making an exponential fit to the forward bias current. Because the result is very sensitive to the voltage range selected, various voltage ranges were used until the correlation factor of the regression was in excess of 0.99. Values lower than this lead to considerable variations in the ideality factor.

With only a continuous density of states, the ideality factor remains about 1.1 but the region across which the device shows an exponential response is much smaller (covering only 0.1 V to 0.3 V). This is illustrated in Figure 3-5.
Figure 3-5: Forward bias I-V characteristics for some the simulations described in Table 3-2. ID#2 is a c-Si p-i-n diode, ID#11 is a diode with a continuous density of states and ID#15 is a diode with an a-Si model.

The shunt resistance and the reverse bias current can be seen to increase. This is indicative of a leakage current flowing across the junction, driven by recombination in the intrinsic layer. Initially surprising is the increase in series resistance. As the mobility has not changed, the series resistance should remain the same, unless the carrier concentration in the intrinsic layer has changed. Figure 3-6 shows how the electron concentration in the diode with a defect density has dropped by nearly two orders of magnitude, explaining the increased series resistance. The drop in carrier density occurs because of the recombination and trapping through the mid gap states.

Note that the electron concentration in the p and n layers varies in the simulations with defect density. This emphasises Hack and Shur's point that the recombination in the contacts is important and should not be neglected. Also, note how the carrier concentration peaks near the p/i interface, again purely an artefact of the recombination model (see Figure 3-6).
Figure 3-6: Plot of electron concentration versus depth for a p-i-n diode with out a defect density (ID#2) and for a p-i-n diode with a defect density (ID#11) given by equation (2-48). The n layer begins at x=0 and goes to x~0.2.

As the mobility is lowered (ID#12,13,14), the same increase in series and shunt resistance is seen as with the earlier simulations (ID#5,6,7). However, this time the ideality factor changes significantly when the electron mobility is reduced. This is an anomalous effect and is not observed when both mobilities are reduced or when only the hole mobility is reduced. The cause of this effect is clearer when the asymmetry in the density of states is examined. With low electron mobility, the majority carriers are now holes. Since there are more defect states closer to the valence band than to the conduction band the capture of holes is more significant. Therefore, the forward bias current is increasingly dominated by recombination of holes (and hence the ideality factor goes to 2). When both carrier mobilities are reduced, the majority carriers are electrons that undergo lower recombination and so the forward bias characteristics are no longer dominated by recombination (and hence the ideality factor goes towards 1).
For the low mobility simulations the shunt resistances, series resistances and reverse bias currents follow the same patterns observed in the earlier simulations (ID#5-8).

In simulation ID#15 the low mobility and continuous defect density of simulation ID#14 are combined with the full set of a-Si parameters given in Table 3-3. These parameters along with the continuous density of states define the basic electrical model that is used in this thesis to describe the a-Si photodiodes.

The observed increase in reverse bias resistance and subsequent reduction in the saturation current, are due to the increased band gap used. This means that there are fewer thermally generated carriers at room temperature (i.e. the carrier density in a-Si is lower than that in c-Si). The increase in ideality factor is also a symptom of this effect, as there are fewer carriers in the bands then the recombination term begins to become increasingly important. To illustrate this note that in equation (2-30) the diffusion term in the diode equation depends on the square of the carrier concentration. However, the recombination term is only linear in carrier concentration. It therefore follows that recombination terms decay slower than diffusion terms and so become dominant at low carrier densities.

The series resistance term given in Table 3-2 is only included for completeness. As can be seen, it has a large value, the reason for this being the large band gap and high recombination. The result is a very low carrier concentration in intrinsic layer, even in forward bias. Figure 3-7 shows how the hole concentration in the intrinsic layer is very low in the intrinsic region of the diode. Clearly, a very large bias must be applied before the diode begins to be limited by the bulk series resistance.
Another item that is of interest is how varying the thickness of the diode affects the DC characteristics (c.f. Section 2.6). In Figure 3-8 below, the IV characteristics for diodes of various thicknesses are shown. All of these diodes use identical models to those used in simulation ID#1 except for the thickness variation. Note how the forward bias series resistance drops (i.e. forward bias current increases) as the diodes become thinner. This is expected, as with a thinner device the bulk resistance must be less. The diode in reverse bias shows the same trend. It is worth noting that from equation (2-37) such dependence in the reverse bias current is to be expected due to the thickness term in the denominator.
Figure 3-8: (a) The forward bias characteristics of diodes of varying thicknesses using c-Si models (Simulation ID#1). (b) Reverse bias characteristics of diodes using the same models.
A similar set of simulations was run for diodes with the a-Si model used in simulation ID#15 and the results are shown in Figure 3-9. Note that in the exponential region the diode shows no dependence on thickness. In addition, the reverse biased diode also shows no dependence on thickness, only the 3 micron diode shows any deviation from the standard plot. This deviation is sufficiently small to suppose that this could be numerical error introduced in the inevitably coarser mesh used in the thicker device. As mentioned in chapter 2, this thickness dependence is often attributed to a charge density at the p/i interface. Here the thickness independence comes about as a product of the recombination through the continuous defect density.

This is illustrated in Figure 3-10, which shows the recombination rate and carrier concentrations through a 1 and 3 micron device in reverse bias. Note how in both devices there is a high level of recombination in the p layer. This is to be expected when the density of traps is greatest near to the valence band, as is true in a-Si. The result is a depletion region at the p/i interface that is independent of the thickness of the device. Thus, as proposed in chapter 2, the thickness independence of the a-Si photodiode is a property of the asymmetry in the density of states. It should not be attributed to the presence of a built in layer of defects.
Figure 3-9: (a) The forward bias characteristics of diodes of varying thicknesses using a-Si models (c.f. Simulation ID#15). (b) Reverse bias characteristics of diodes using the same models.
Figure 3-10: A plot of the simulated carrier concentrations and recombination rates through in a-Si photodiodes of varying thicknesses. (a) is a diode with a thickness of 1 µm and (b) has a thickness of 3 µm, other than that all parameters are identical to simulation ID#15.
Simulation of Photodiode Structures

### Basic Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value for e-Si</th>
<th>Value for a-Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band Gap (eV)</td>
<td>1.08</td>
<td>1.72</td>
</tr>
<tr>
<td>Mobilities (cm² s⁻¹ V⁻¹)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrons</td>
<td>1000</td>
<td>10⁷</td>
</tr>
<tr>
<td>Holes</td>
<td>500</td>
<td>1⁷</td>
</tr>
<tr>
<td>Thermal Velocities (cm² s⁻¹)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrons</td>
<td>1.03 x 10⁷</td>
<td>1 x 10⁷</td>
</tr>
<tr>
<td>Holes</td>
<td>1.03 x 10⁷</td>
<td>1 x 10⁷</td>
</tr>
<tr>
<td>Conduction Band DOS (cm⁻³)</td>
<td>2.8 x 10¹⁹</td>
<td>1.0 x 10²²</td>
</tr>
<tr>
<td>Valance Band DOS (cm⁻³)</td>
<td>1.04 x 10¹⁹</td>
<td>1.0 x 10²²</td>
</tr>
</tbody>
</table>

*These values are not used exclusively for a-Si models. They are used in any low mobility simulation and any simulation for a full a-Si model. The basic simulations use the values for c-Si.

### Constants

<table>
<thead>
<tr>
<th>Constant</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boltzmann’s Constant (kₙ)</td>
<td>1.38066 x 10⁻²³ J/K</td>
</tr>
<tr>
<td>Elementary Charge (e)</td>
<td>1.6023 x 10⁻¹⁹ C</td>
</tr>
<tr>
<td>Permittivity in vacuum (ε₀)</td>
<td>8.85418 x 10⁻¹⁴ F/cm</td>
</tr>
<tr>
<td>Thermal Voltage (kₙ T)</td>
<td>0.0258502 V</td>
</tr>
</tbody>
</table>

### Recombination Models

Direct Optical and Auger recombination are never used.

Simulations using SRH recombination assume recombination lifetimes of \( τ_n = 0.1 \) μs and \( τ_p = 0.1 \) μs.

For those models using a continuous defect distribution, the defect density for acceptors and holes was defined by equation (2-16) and the recombination is calculated using equation (2-18) with the following parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{min} = \sigma_{Amin} )</td>
<td>1 x 10¹⁶ cm⁻³</td>
</tr>
<tr>
<td>( E_{nv} )</td>
<td>1.07 eV</td>
</tr>
<tr>
<td>( E_d )</td>
<td>0.088 eV</td>
</tr>
<tr>
<td>( E_h )</td>
<td>0.053 eV</td>
</tr>
<tr>
<td>( \sigma_A )</td>
<td>1 x 10⁻¹⁶</td>
</tr>
<tr>
<td>( \sigma_{A} )</td>
<td>1 x 10⁻¹⁴</td>
</tr>
<tr>
<td>( \sigma_{D} )</td>
<td>1 x 10⁻¹⁴</td>
</tr>
<tr>
<td>( \sigma_{D} )</td>
<td>1 x 10⁻¹⁶</td>
</tr>
</tbody>
</table>

Here \( \sigma \) represents the cross section for interaction, a subscript e marks a process involving an electron and h indicates a process involving a hole. Superscript A represents acceptor states whilst D represents donor states.

Table 3-3: Comparison of the default values used in c-Si simulations and a-Si simulations.
3.2.6 Photodiode Response

As discussed in chapter 2, the principle figure of merit for a photodiode is the ratio of the photocurrent to the dark current. In this set of simulations a photodiode was illuminated with varying intensities of light and the signal to noise ratio was determined. For the first set of simulations, the simple p-i-n diode with c-Si parameters was used and the optical absorption was changed from 1 to $10^4 \text{ cm}^{-1}$ (a value typical for a-Si in green light, i.e. at 550 nm). The diode was reverse biased at 5V, since this bias voltage allowed the data for reverse bias current from Table 3-2 to be used.

The results are shown as a log-log plot, in Figure 3-11. Here the signal has been taken as the ratio of the photocurrent to the dark current plus 1, thus normalising zero signal as 1. In addition, the bold line represents the expected response if the photodiode was an ideal linear detector. It can be seen that the photocurrent increases linearly with intensity as would be expected. This also illustrates that higher optical absorption means that the diode can respond to lower levels of illumination.
The next step was to introduce a low mobility model, the same one used in simulation ID#7. Here only two values for optical absorption were used, $10^3$ cm$^{-1}$ and $10^4$ cm$^{-1}$. The results (in Figure 3-12) show a similar response as in the case of c-Si. However, the signal is higher, this is because the lower mobility gives a lower dark current. Usually, the lower mobility would enhance trapping and the signal would degrade. However, in this case the recombination model was turned off and so we see this increase in signal.
Finally, Figure 3-13 shows what happens when the a-Si models are introduced. At first sight, these may seem to be identical in form to the above simulations. However note that as the intensity starts to reach high values (> $1 \times 10^5$) the signal is no longer linear with intensity. Here the generation of carriers is so high that the device starts to exhibit bimolecular recombination\(^4\). At high intensities, the quasi-fermi levels are close to the band tails. Therefore, the density of deep traps has effectively increased. This lowers the effective carrier lifetime and hence the photoconductivity.

Because of the complications that it produces, future simulations will avoid the bimolecular regime. This is done by running the simulations at an intensity of 1 W/cm\(^2\), a value that gives a good signal whilst maintaining linearity.
3.2.7 The Influence of Defect Density

Though we assume the density of states that is described by equation (2-48), it is worth noting how changes to this density of states influences the operation of the photodiode. A series of simulations was run that varied the parameters that define this density of states, $g_{dmin}$, $E_a$, $E_d$ and $E_{mv}$ (the default values are shown in Table 3-3 and these values are explained in chapter 2).

The first set of simulations varied the defect density from $1 \times 10^{16}$ cm$^{-3}$ to $1 \times 10^{17}$ cm$^{-3}$. A comparison of the forward bias and reverse bias characteristics are shown in Figure 3-14. It can be seen that with increasing defect density the recombination goes up. Hence, there is an increase in the reverse bias current and the forward bias series resistance.

Figure 3-13: A log-log plot of normalised signal versus incident light intensity for devices simulated using a-Si models. Note the non-linear region that appears beyond $10^5$ W/cm$^2$. 
Figure 3-14: A comparison of the current-voltage characteristics for a-Si diodes with defect densities of $1 \times 10^{16}$ cm$^{-3}$ (Base Line) and $1 \times 10^{17}$ cm$^{-3}$. The forward bias characteristics are shown in (a) and the reverse bias characteristics are shown in (b).
The change in the photoresponse of the diode was also investigated and this is illustrated in Figure 3-15. Broadly the photoresponse is the same, however, increasing defect density does mean a reduction in photocurrent and the non-linearities at higher defect density become increasingly worse. This is in line with the origin of the high intensity non-linearities being due to bimolecular recombination.

![Figure 3-15: The current through photodiodes with varying defect density, versus the intensity of illumination.](image)

The next parameters of interest are $E_a$ and $E_d$, which are the slopes of the band tails. With a steeper slope (i.e. a smaller value of $E_a/E_d$), the defect density at the band tails goes up ($E_d$ pushing up the donor density at the conduction band and $E_a$ pushing up the acceptor density at the valence band). Figure 3-16 shows the forward and reverse bias characteristics for simulations in which $E_a$ and $E_d$ have been varied independently. The transition from $E_a=0.1$ to $E_a=0.053$ has little effect on the IV characteristics, however $E_a=0.01$ gives a drastically lower reverse bias current and a
much larger forward bias current. This is because near the valence band edge the density of acceptor states is so high that the majority of the holes are trapped there. However, these states are too close to the valence band to act as efficient recombination centres and so the recombination drops to zero.

For the case of reducing $E_d$, the number of traps near the band tails increases. This has an adverse effect on electron mobility (hence the reduction in the forward bias current) and increases the recombination rate (hence the increase in reverse bias current).

![Figure 3-16: Forward and Reverse bias characteristics for a-Si photodiodes with different density of states. (a) shows the forward bias characteristics for devices where $E_a=0.1$, 0.053 and 0.01, (b) is the reverse bias characteristics for these devices, (c) and (d) present, respectively, the forward and reverse bias characteristics for diode with $E_d=0.1$, 0.088 and 0.01.](image-url)
Considering how the photoresponse varies with changes in $E_a$ and $E_d$, it is expected that increasing $E_d/E_a$ should improve the linearity of the photoresponse due to the reduction in tail states. Figure 3-17, illustrates that this is in fact so. The very non-linear shape for the case when $E_d=0.01$ and $E_a=0.01$ is due to the exceptionally high defect density at the band edges that this value leads to.

The final parameter that can be changed in the model's density of states is the location of the minima in the density of states, $E_{mv}$. The results of simulations with values of $E_{mv}=1.2, 1.07$ and 0.8 eV are given below in Figure 3-18 and Figure 3-19. From these figures, it is clear that the location of the minima in the density of states plays an important role in determining the diode characteristics. If $E_{mv}$ is greater than half the band gap (0.86 eV in this case), then the majority of traps are acceptor like. The result is high recombination through acceptor states, hence a high reverse bias current and a low forward bias current. When $E_{mv}=0.8$ the device begins a transition towards having a greater density of donor like traps. This would result in the majority carrier being holes. However, because the minimum is close to the middle of the mobility gap, recombination is minimised and the effective series resistance is also minimised. The result is a much large forward bias current and lower reverse bias current.
Figure 3-17: Current as a function of intensity. Here (a) shows the results of simulations where $E_a$ has been given values of 0.1, 0.053 and 0.01, (b) shows the same information where $E_d$ has been given values of 0.1, 0.088 and 0.01.
Figure 3-18: (a) illustrates the effects of varying $E_{mv}$ on the forward bias characteristics of an a-Si diode, (b) shows how the reverse bias characteristics are influenced.

Figure 3-19: The influence of $E_{mv}$ on the response of a photodiode to illumination of varying intensities.

**Section 3.3 - The Basic Photodiode Sensor**

It is of interest to examine how a photodiode with different cathode and anode sizes will respond to incident illumination. To do this a simulation is run in which a beam, 2 $\mu$m in width, is incident on the cathode side of the photodiode with the structure
shown in Figure 3-2. The beam is brought in perpendicular to the front surface of the photodiode and reflections and absorption from the front and back contacts are neglected.

The device is then held at 5 V, reverse bias, and the current measured. By moving the position of the incident beam across the surface of the device, the position response of the photodiode can be assessed.

To speed up simulation times, we note that because of symmetry it only necessary to examine the response across the part of the diode illustrated in Figure 3-2. Figure 3-20 shows this by plotting the anode current at each end of the incident beam’s width, for beams at various positions. The diamond markers represent the start and finish of the adjacent beams, and the line represents the current measured when this beam was illuminating the device with an intensity of 1 W/cm². The optical absorption coefficient is assumed to be $10^4$ cm$^{-1}$, and the diode is reverse biased at 5 V.

It can be clearly seen that at the edge of the anode contact the current drops down, rapidly falling to the level of the dark current. This means that no signal would be obtained from this region of the photodiode. There is a poorly resolved region at the edge of the photodiode where the photocurrent is still slightly above the dark level. Note that in this region the doping is a constant $1 \times 10^{15}$ cm$^3$. Once the 2 micron region is bypassed then the diode fails to respond.
Figure 3-20: A plot of the current in a photodiode (illustrated in Figure 3-2) versus the lateral position of an incident beam of light with an optical power of 1 W/cm².

This means therefore that this type of photodiode is worse than a diode that is 60 microns wide and isolated by a trench at each edge. That this is so is not immediately obvious, as it would seem that both diodes would have equivalent performance. However, this ignores the issue of isolation. Consider Figure 3-21, here we can see that tracks underlying the photodiode could potentially act as gates. Such tracks may be the data and gate lines. If they become negatively biased then this TFT could switch on and significant conduction could occur between adjacent diodes. Theoretically, this should not happen as the relative biases along the gate and data lines will positive for an n type TFT. However, the possibility of such parasitic action does exist and is an important consideration.
It is of interest to examine what happens when the anode and cathode are swapped. The reason for this is the asymmetry in the carrier concentrations shown in Figure 3-10. The high charge concentration at the p/i interface causes an increased electric field. If the anode and cathode are swapped then the majority of charge will be generated at this interface. The result should be an improvement in charge collection. Another simulation was run as described earlier in this section, but this time with a 100 μm anode as the top contact and a 60 μm cathode as the bottom contact. The results are presented in Figure 3-22. Though similar, this photodiode shows some improvements over the previous diode. Across the region of the Anode, the photocurrent is slightly higher. At the contact edges, the current does not drop as rapidly. More importantly the current does not drop down to the dark current level of ~1x10^{-17} A/μm. Instead, it falls off slowly to a value of approximately 1x10^{-14} A/μm,
and represents a photocurrent created by the collection of carriers at the p/i interface. This is an improvement over a photodiode structure using isolation, however this must be placed in context. With a typical electrode dimension of 30 µm for this region the total current developed by a 1W/cm² beam is only 0.3 pA, corresponding to a responsivity of 29 µA/W. This should be compared with an ideal diode value of 0.44 A/W and the value of 20 mA/W for the photodiode in the region of the anode and cathode.

![Anode Current vs Lateral Beam Position](image)

Figure 3-22: A plot of the current in a photodiode versus the lateral position of an incident beam of light with an optical power of 1 W/cm². The structure is similar to that given in Figure 3-2, but here the anode and cathode have been interchanged.

Section 3.4 - The Focussed Photodiode

Given the observation in Section 3.3 that a considerable amount of signal is lost in regions of the diode where part of the contact is missing, can this signal be recovered?
This section uses the simulations to explore structures that can perhaps reduce the signal lost in these regions of the photodiode. The common factor for all of these structures is that they attempt to introduce a lateral electric field to allow collection of charge from regions of the diode beyond the edges of the contacts.

3.4.1 The Junction Focussed Photodiode

Figure 3-23 shows the structure for this photodiode. The principle is that an extra focussing electrode be placed in the dead space between diodes. This should result in an electric field that is capable of driving carriers towards the signal electrode (the anode in this case). In addition, the presence of the focus electrode acts as an isolation to prevent the parasitic FET effect mentioned in section 3.3. For these simulations, the doping at the focus electrode is the same as that in the cathode (1x10^{15} cm^{-3} phosphorous).

The flaw with this structure however lies with the vertical n-i-n structure that is created. If the depletion regions formed by the n regions should meet then a diode action takes place and the current will rise exponentially until limited by the build up of space charge in the device. The point at which this occurs is called reach through and the voltage at which it occurs is the reach through voltage.
Simulation of Photodiode Structures

Figure 3-23: A schematic for a focused junction photodiode. The parameters are the same as for the basic photodiode (Table 3-1) but with the extra parameters used to define the focus electrode given in (Table 3-4) below. The doping in the focus electrode is always taken to be the same as that in the top electrode.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Anode - Focus Spacing</td>
<td>15 μm</td>
</tr>
<tr>
<td>12</td>
<td>Focus Lateral Doping Overlap</td>
<td>2.0 μm</td>
</tr>
<tr>
<td>13</td>
<td>Focus Vertical Doping Width</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>14</td>
<td>Focus Doping Tail</td>
<td>0.1 μm</td>
</tr>
</tbody>
</table>

Table 3-4: Additional parameters for the definition of a focused junction photodiode as shown in Figure 3-23, above.

This action is illustrated in Figure 3-24 below. Here the current flowing through the focus electrode and the anode is plotted as the cathode bias is varied. The focus electrode is fixed at 2.5 V; this means that when the focus and cathode voltages are equal the device is in reverse bias. The cathode current is not plotted, as it is always
approximately equal to the larger of the two currents flowing. Note how when the cathode voltage and focus voltage match the focus current drops to a near zero value and the current flowing through the device is equal to that leaking through the p-i-n diode. The narrow voltage range over which this occurs indicates that the reach through voltage is only a few hundred millivolts. Clearly, the only reasonable range to operate this device in is with the cathode voltage equal to the focus voltage. It has been found in a series of simulations, that if left floating the focus electrode will rise to the potential at which the anode and focus currents are equal. This suggests the possibility that the focus electrode need not be biased but may be left to rise to its own value. Such a structure would immensely simplify processing and operation of such a diode when in an active matrix. However, there are two reasons why this approach has not been taken in the series of simulations. Firstly simulations with floating boundary conditions are slow to converge and hence computationally intensive. Secondly, without a path to earth it is possible that illumination over the junction photodiode would generate an immovable space charge. This would be difficult to observe in the simulations as it would tend to cause divergence and the effect would be indistinguishable from numerical errors and poor gridding.
The only way in which to increase this reach through voltage is to either increase the diode thickness or to reduce the doping in the n layers. This first option is not feasible, as the growth of thicker layers of a-Si is plagued with problems. A reasonably yielding process is certainly out of the question given current technologies. Reducing the doping in the n layers would lead to increased leakage across the junctions. This is not desirable for a low noise diode.

Having determined that there is only one feasible operating condition; it is now worth measuring the photoresponse of the diode at this operating condition. The same technique that was applied in Section 3.3 has also been applied with the focus electrode and cathode held at 5V (i.e. reverse biasing the diode). However, in this case the beam width was only 1 μm. The results of these simulations are plotted below in Figure 3-25.
Figure 3-25: A plot of the Anode current in a p-i-n photodiode with the addition of a lateral focusing electrode, as described in the text.

Note that with this structure there is negligible improvement at the edge of the anode contact. However, there is some interesting behaviour approaching the focus contact where the anode current begins to fall. If the cathode and anode currents are also plotted, as they are in Figure 3-26, then it can be seen that in this region the signal is collected by the cathode and focus electrodes. Arguably, this signal could be utilised by taking the signal from the cathode electrode instead of the anode. However, note that the gain is only slight. Also, consider that this signal is effectively from the n-i-n diode. Hence, it would be non-linear and noisy. Hence, we neglect this signal as it is of little practical use.
Figure 3-26: Plot of Cathode and Focus Current as a function of the lateral position of an incident beam.

Figure 3-27 shows the results of the same set of simulations but this time with a n-i-p device. Note that as the anode doping has changed to n type then the focus electrode doping has to change to p type. As with the simulations in the previous section the n-i-p photodiode has a superior performance. This is again because of the enhanced field at the p/i top interface where the majority of the optical absorption occurs. Like the p-i-n diode the n-i-p diode sees an enhanced signal in the top contact (this time the anode) but a degradation in the cathode (i.e. bottom) contact. For the reasons discussed in the p-i-n case this signal is also of little practical use.
Unfortunately, there is no observation of a focussing effect, i.e. carriers being repelled by the field of the focussing electrode into the signal electrode. This is primarily because of the geometry of the device. The vertical dimension of the transistor is 1 micron, the horizontal spacing is 15 microns. This means that there is no depletion region formed between the anode and the focus electrode. The result is a narrow depletion region at the edge of the signal electrode of a few microns from which an active signal can be obtained. The presence of the focus electrode is neutralised by the relative distance from the top contact to the focus versus the focus to the bottom contact, the former being smaller than the latter.

In conclusion, besides isolating the device properly the junction focussed photodiode has little to recommend it versus the basic photodiode described in Section 3.3. The
extra processing steps required to manufacture such a photodiode probably do not justify the improvements.

3.4.2 The Insulated Focus Photodiode

In the previous section, it has been shown that the junction focussed photodiode offers little improvement in signal compared to the basic photodiode. One problem with these devices is the photocurrent from the n-i-n or p-i-p diode. When illuminated this diode begins to leak and causes signal loss.

A simple way to overcome this problem is to replace the junction with an insulating layer. This allows any field to be applied to the focus without inducing leakage. The upper limit on the applied field is the dielectric breakdown of the insulator layer, for Silicon Nitrides and Oxides this lies at around $10^7$ V/cm. This gives a maximum voltage of 100 V over a 0.1 μm thick layer. However in practice the insulator will start to conduct significantly at lower voltages (~30-40 V) due to Poole-Frenkel emission or Fowler-Nordheim tunnelling. Because of this, and restrictions in simulation time we will restrict ourselves to voltages of up to 10 V on the focus electrode.

Practically this introduces several complications into the design. One example is the poor quality of any insulating layer that will be exposed to the plasma when depositing the a-Si layers. This means that the presence of trapped charge at the interface, or leakage across the insulator is a possibility. However, for these simulations we simply replace the focus junction with an ideal insulating layer, as illustrated below in Figure 3-28.
In reality, the insulating layer is likely to be PECVD amorphous Silicon Nitride, and so would be deposited as a conformal layer. A dry etch would then be used to pattern the Nitride, leading to a smooth rounded edge, like a spacer in a MOSFET. However, for simplicity it is assumed that the nitride takes on the artificial shape shown above and that its thickness is constant. Ultimately this will not affect the general performance of the device. However, note that with a 0.1 \( \mu \text{m} \) insulating layer the effective distance between the active area of the focus and the anode contact has increased by a few micrometers.
Simulation of Photodiode Structures

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Anode - Focus Spacing</td>
<td>15 μm</td>
</tr>
<tr>
<td>12</td>
<td>Lateral Insulator Thickness</td>
<td>0.2 μm</td>
</tr>
<tr>
<td>13</td>
<td>Vertical Insulator Thickness</td>
<td>0.1 μm</td>
</tr>
</tbody>
</table>

Table 3-5: Additional parameters for the definition of a focussed insulating photodiode as shown in Figure 3-23, above.

It is found that the bias on the focus electrode has no influence on the dark IV characteristics of this photodiode. Therefore, any bias may be applied to the insulated focus. Figure 3-29 below shows the profiles for the net carrier concentration (defined as the modulus of n-p), taken along a cross section taken through the cathode and the insulated focus electrodes. There is a bias of 0V on the cathode and the anode and various biases are applied to the focus electrode. In running these simulations, there are two possible options. Firstly, to isolate devices we wish to deplete the semiconductor of carriers, hence it can be seen that an applied voltage of -5 V (with respect to the cathode voltage of 0V) is the preferred operating condition. Secondly, there is a wish to drive carriers from the focus contact to the signal contact. To do this the focus contact bias must be greater than the bias on the top contact. Obviously, these two aims are contradictory so both devices have been simulated.
Figure 3-29: A plot of the absolute net carrier concentration versus depth. This is taken as a cross section from an insulated focus device (see Figure 3-28) in the region above the focus electrode. The cathode lies at x=0 and the insulator above the focus starts at x=8.0. The focus electrode is ranged from -10 to +10 Volts whilst all other voltages are kept constant.

Figure 3-30 below shows the results of a response simulation on a p-i-n insulated focussed photodiode. Here the Anode current is plotted versus the beam position for focus biases of +10V and -5V with respect to a cathode bias of 5V and an anode bias of 0V. Again there is no appreciable influence from the focus electrode; even elevated biases do not allow sufficient lateral field to be developed for any type of focussing effect to be observed. In fact, the response is worse than the junction focussed photodiode and the basic photodiode as the photocurrent drops rapidly at the junction edge.
Figure 3-30: A plot of Anode currents for a p-i-n insulated focussed photodiode with -5 V and +10 V on the focus electrode.

The results for an n-i-p diode are given below in Figure 3-31; here the first signs that a focussing effect may be possible can be seen. When the diode is biased so that the focus electrode has a larger bias than the cathode electrode, the photocurrent obtained is larger. This could indicate that the photogenerated charge is pushed towards the signal (cathode) electrode by the biasing electrode. Though there is an improvement in characteristics with focus bias the results still fall short of those obtained from the junction focussed photodiode because of the lower collection at the cathode contact edge.
3.4.3 The Skewed Photodiode

Though the insulated focussed photodiode is capable of introducing a limited lateral field, it is limited by the geometry. The vertical distance between contacts is far less than the horizontal region. This restricts the operating voltage and prevents a focussing field from taking full effect. A proposed geometry to overcome this is the skewed photodiode (see Figure 3-32). Note that this is the first structure where we actually show two adjacent pixels. This is necessary to illustrate the overlaps between pixels.

In this device, the top contact is also patterned but it is longer and shifted slightly. Because of this, there must be a lateral field present between one edge of the cathode
and the anode beneath it. This has several drawbacks. Firstly, there is a loss of active area due to the smaller cathode area. It is hoped that this may be recovered by the lateral field’s introduction. A second disadvantage is that there is now a capacitive coupling between the top cathode and the adjacent signal electrode, giving rise to the potential for more crosstalk. However, there are several advantages as well. There is a lower capacitance due to the reduced contact area, and we gain the ability to tailor the lateral electric field. Also, there is less loss of light at the photodiode interface due to reflections at the top contact.

Figure 3-32: A schematic of two adjacent skewed photodiodes. The numbers mark the parameters defined below in Table 3-6.
Simulation of Photodiode Structures

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cathode Metal Thickness</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>2</td>
<td>Anode Metal Thickness</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>3</td>
<td>Diode Thickness</td>
<td>1.0 μm</td>
</tr>
<tr>
<td>4</td>
<td>Anode Width</td>
<td>60 μm</td>
</tr>
<tr>
<td>5</td>
<td>Device Width</td>
<td>100 μm</td>
</tr>
<tr>
<td>6</td>
<td>Cathode Width</td>
<td>80 μm</td>
</tr>
<tr>
<td>7</td>
<td>Cathode Doping Overlap</td>
<td>2 μm</td>
</tr>
<tr>
<td>8</td>
<td>Anode Doping Overlap</td>
<td>2 μm</td>
</tr>
<tr>
<td>9</td>
<td>Cathode-Anode Overhang</td>
<td>25 μm</td>
</tr>
<tr>
<td>10</td>
<td>Cathode Doping Tail</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>11</td>
<td>Cathode Doping Thickness</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>12</td>
<td>Anode Doping Tail</td>
<td>0.1 μm</td>
</tr>
<tr>
<td>13</td>
<td>Cathode Doping Thickness</td>
<td>0.1 μm</td>
</tr>
</tbody>
</table>

Table 3-6: Parameters used to define the skewed photodiode used in the simulations. The numbers refer to the lengths marked on Figure 3-32.

This device has been simulated; however, to speed up simulations the device has only been simulated from the midpoint of the base electrode to the midpoint of the adjacent base electrode. This makes for a smaller device size without losing the information required for the simulation. Figure 3-33 and Figure 3-34 show how the anode and cathode electrodes in a p-i-n diode respond to the position of the incident beam. To the left of the Right anode contact, an increase in current is observed which indicates a successful improvement in the collection from this part of the device. This increase is however mirrored in both cathode currents, why the left cathode current should suddenly surge is not immediately obvious. The fact that this increase is almost coincident with the right anode however does suggest an answer. It seems probable that the anode is acting as the gate of a JFET transistor. Without illumination, the JFET is off and there is no conduction. However, the photogenerated carriers switch
on the JFET causing carriers to flow directly from cathode to cathode. The same effect is presumed to be responsible for the anode current.

Figure 3-33: A plot of the anode currents, as a function of the lateral beam position for a p-i-n skewed photodiode.
Figure 3-34: A plot of the cathode currents, as a function of the lateral beam position for a p-i-n skewed photodiode.

The n-i-p structure shows similar behaviour, because of the larger mobility of electrons it is again the n-type cathodes that are affected by this parasitic JFET. This can be seen in Figure 3-35 and Figure 3-36. The effect is more pronounced this time, because the region covered by the anode is larger.
Figure 3-35: A plot of the anode currents, as a function of the lateral beam position for a n-i-p skewed photodiode.

Figure 3-36: A plot of the cathode currents, as a function of the lateral beam position for a n-i-p skewed photodiode.
The benefits of the extra signal from this device are dubious. Though there are signs of an improved response in the device, the JFET effect is not necessarily beneficial. It would seem that there would be a strong non-linearity in the effect. This is because the exact location of charge generation will vary according to the wavelength of the incident light. Hence, the amount of injected charge switching the parasitic JFET will vary, as will the final signal.

3.4.4 The Insulated Focus Skewed Photodiode

From the previous work on the skewed photodiode and the focussed photodiode the next logical step is to try and integrated the two devices. Figure 3-37 illustrates how insulated focus contacts could be added to the skewed photodiode structure of Figure 3-32. For this structure, the dimensions specified for the skewed photodiode still hold. The insulated contacts are introduced as 10 μm wide aluminium strips that lie half way between adjacent cathode (or anode) contacts. This strip is then isolated from the silicon in the vertical direction by a 0.1 μm thick silicon nitride insulating layer. In the lateral direction, the silicon nitride is 1 micron thick.
Figure 3-37: An illustration of the inclusion of insulated focussed contacts into the skewed photodiode.

As with the insulating photodiode, the device may be operated in two regions. The bias conditions for these are illustrated in Figure 3-38. The first is a focussing mode where the focus electrodes are biased so those electrodes adjacent to anodes have positive applied voltages greater than that applied to the cathode. Electrodes adjacent to cathodes are set to have negative voltages equal to those applied to the focuses adjacent to the anodes. This should have the next effect of driving carriers towards the signal electrodes. In depletion mode, the voltages are applied in the different sense to pinch off any conducting channel that may form between adjacent electrodes.
Figure 3-38: Illustration of the biasing of insulated focus skewed photodiodes. The diode to the left is in focus mode whilst the diode to the right is in depletion mode.

The results for a p-i-n photodiode in depletion mode are given in Figure 3-39 and the results for focus mode are given in Figure 3-40. It can be seen that the results are identical to those for a p-i-n skewed photodiode without the addition of the insulating contacts.
Figure 3-39: The photoresponse of a p-i-n insulated skewed, focussed photodiode when operated in depletion mode. (a) shows the anode currents in the device and (b) gives the cathode currents. In this device, the base insulated contact runs from 45 to 55 μm and the top insulated contact goes from 60 to 70 μm.
Figure 3-40: The photoresponse of a p-i-n insulated skewed, focussed photodiode when operated in depletion mode. (a) shows the anode currents in the device and (b) gives the cathode currents. In this device, the base insulated contact runs from 45 to 55 \( \mu m \) and the top insulated contact goes from 60 to 70 \( \mu m \).
Figure 3-41: The photoresponse of a n-i-p insulated skewed, focussed photodiode when operated in depletion mode. (a) shows the anode currents in the device and (b) gives the cathode currents. In this device, the base insulated contact runs from 45 to 55 µm and the top insulated contact goes from 60 to 70 µm.
Figure 3-42: The photoresponse of a n-i-p insulated skewed, focussed photodiode when operated in focus mode. (a) shows the anode currents in the device and (b) gives the cathode currents. In this device, the base insulated contact runs from 45 to 55 µm and the top insulated contact goes from 60 to 70 µm.
Figure 3-41 and Figure 3-42 have similar response curves to the n-i-p skewed photodiode. However, when operated in depletion mode (Figure 3-42) the effect due to the parasitic JFET is suppressed over a wider region. This is presumably because the insulated contact has depleted the conducting channel between the two n type contacts preventing conducting. Hence, even when the JFET is switched no my in the coming beam of light the FET counters this action and no extra current flows.

**Section 3.5 - Conclusions**

In this chapter, we have developed a model for a-Si diodes. The various components of this model have been examined individually to assess the impact that each component has. This has also allowed the various models to be checked for anomalies. It was found, as expected that the diode ideality factor is close to 1 for a-Si diodes, and is unaffected by changes in contact geometry or mobility. However, as expected it does vary when recombination is introduced. When a recombination model is in use, be it a single level Shockley-Read-Hall model, or a continuous defect model, the ideality factor does depend on the mobility. This dependency occurs because neither the recombination nor the diffusion currents dominate. Increasing or decreasing the mobility increases or decreases the relative contribution of the diffusion currents.

The forward bias resistance was found to vary with mobility as expected. In addition, the reduction of mobility and the addition of a continuous density of states increased the resistance as would be expected, as a-Si is far more resistive than c-Si.

The saturation current in reverse bias was found to increase with the introduction of recombination models. However, the increase in the mobility gap of a-Si, with respect
to the band gap of c-Si means that a-Si diodes show a much lower reverse bias current.

It was shown that the introduction of an asymmetric density of states is sufficient to account for the thickness independence of the current-voltage characteristics of a-Si diodes. This is due to the high recombination at the p/i interface.

In terms of the response to light, it was shown that as expected, a c-Si photodiode responds linearly with intensity over a large range of values. The upper and lower limits being set by the onset of high carrier concentration effects (not modelled here) and the dark current respectively. It was illustrated that the introduction of the a-Si model introduces a non-linearity into this signal. However, there is an acceptable operating region in which the response is nearly linear.

The current voltage characteristics of a-Si diodes were examined for different defect densities. It was found that increasing the number of defects would increase the reverse bias current and lower the forward bias current of diodes, consistent with increased recombination. This higher defect density also correlates with an increase in the non-linearity of the photodiode’s response with changes in the intensity of incident light.

Modifications to the slopes of the band tail defects where found to change the current voltage characteristics in a similar manner. This was for the same reason that increasing the sloped of the band tails increases the total defect density.

Changing the location of the minima in the defect density with respect to the band tails was found to modify the properties of the diode. When the defect density lies closer to the valence band, recombination falls and diode characteristics improve. However,
Simulation of Photodiode Structures

as the minima in the defect density shifts towards the conduction band, recombination increases. This behaviour equates with an increase in the availability of recombination centres, though there is a consequent reduction in trapping sites.

All of these conclusions were shown to be in agreement with the internal values for the simulator such as carrier concentration and electrical potential.

These models where then applied to various diode structures to see if it was possible to engineer a lateral drift field to improve the active area of the photodiodes. In all cases, any increase was found insignificant. The principle reason for this was that the vertical dimension compares to the lateral dimension by a ratio of about 1:15. Hence, lateral fields are generally lower than the vertical fields. However, interesting behaviour was observed in all of the simulations.

The first device, the basic photodiode, with a shortened contact on the base of the device, showed good charge collection over the contacts. The simulations indicated that if a n-i-p structure were to be used then there would be collection from the regions beyond the shortened contacts. However, the magnitude of this signal would be very small.

The introduction of a junction to act as a focus electrode was found to reduce the current from the shorter, signal, electrode slightly. An increased response was noted over the areas of the junction, but this was dismissed as being unusable signal because of the inherent non-linearities in space charge diodes and the difficulties of making connections to the contacts.

Insulating the junction removed these problems and allowed the possibility of focusing the signal or of improving isolation between devices. Improvement was
noticeable when using a diode in focussing mode compared to when being used in isolation mode. Evidence for a focussing effect perhaps, unfortunately there was no overall improvement compared to the basic photodiode.

The skewed photodiode was turned to as a means of introducing a larger lateral field by offsetting the electrodes on the diode. Examining the diode revealed an increase in signal in regions of the diode not covered by both electrodes. However, when analysed further it became apparent that this signal was possibly due to a phototransistor effect. This would be potentially non-linear and again of little use. The addition of insulating contacts showed that this effect could be switched off, but still the signal was no better than that from an ordinary photodiode.

The conclusion from this is that the additional processing steps involved in manufacturing these novel structures do not justify the very small gains that may be had from using these devices.

Section 3.6 - References


Chapter 4  Dangling Bond Recombination in a-Si:H

Section 4.1 - Introduction

The problem of dangling bond recombination was first addressed in Chapter 2 where a brief review of approaches to the problem was given. It was also noted in the previous chapter that dangling bond recombination must be accounted for, to give an accurate simulation of device characteristics in a-Si:H. This is especially important for determining the signal to noise ratio. There are also reasons to believe that issues such as device lifetime and degradation can only be addressed by fully considering the structure of the dangling bond defects.

However, these issues have been ignored thus far and it is tempting to use the model of Hubin et al.\textsuperscript{3} within the Silvaco simulators. Unfortunately, because this model only applies at high generation rates, it cannot be used in our case. This is because, to ensure convergence, the simulator must be taken from a zero bias, zero illumination solution slowly to a biased and illuminated solution. Using the Hubin et al. model creates excessive generation of carriers when there is no illumination and the simulators fail to converge.

Currently, there are two dominant models to describe dangling bonds in a-Si, both taking quite different approaches. In this chapter, we review these models with the aim of incorporating them into our simulations. As will be shown there are several issues that prevent this from being done. We also outline an approach that was originally thought to be promising, but as shall be shown, is unusable.
First we consider the nature of the defects in a-Si and define the defect density of states. After this, two models are discussed which offer different approaches to the issue of defects in a-Si materials. The first of these models is the defect pool model, which allows the calculation of a defect density from thermodynamic considerations. However, this approach does not readily produce the recombination function, nor does it produce a consistent free carrier concentration. The second model is the rate equation approach, which lacks a first principles derivation of the defect density but does have the advantage of producing the recombination function. Finally, we close this chapter with a discussion of the transient problem.

**Section 4.2 - The Nature of Defects in a-Si**

In Chapter 2 we touched briefly upon the mobility gap and localised states in amorphous materials. It is worth considering the physical origin of such defects before studying some physical models that we could use to describe them.

In effect there are two classes of defects in amorphous materials, the first arise due to the random nature of the lattice and the second arise due to defects in the lattice. The latter are commonly referred to as dangling bonds. Defects that arise from the random nature of the lattice (also known as topological defects) are difficult to understand physically but we can picture these defects as arising from potential fluctuations. These fluctuations arise throughout the lattice because of the inherent randomness of the lattice. Figure 4-1 illustrates how such a fluctuation can lead to a localised state for a simple 1D case (see also Elliot's description of localisation for an analytical treatment).
It is worth noting that these defects tend to be occupied by single electrons. To place two electrons into a single state would probably require binding energies so large that having both electrons in the conduction band is more energetically favourable. These defects are frequently associated with the localised states at the band edges. The result in the so called 'band tails'.

Figure 4-1: Illustration of how a randomly varying potential can contain fluctuations that will act as trapping centres for electrons.

Dangling bond defects arise when an atom cannot bind to any of its neighbours. The result is a lone electron in a binding orbital, which has an energy that usually places it in the middle of the a-Si mobility gap. In some cases, it can be energetically favourable for the lone electron to escape the site or to capture a second electron, giving rise to amphoteric dangling bonds. As has already been discussed in Chapter 2, these bonds have only recently been recognised as playing a major role in describing the reverse bias currents of a-Si photodiodes.

These defects are even more interesting when the role of weak bonds and hydrogen is considered. Hydrogen acts to passivate these dangling bonds, reducing their numbers,
but it can move and diffuse through the lattice leading to a changing defect bond density. Also hydrogen 'scavenges' weak Si-Si bonds and will sometimes break such bonds creating a Si-H bond and an additional dangling bond. The Staebler-Wronski effect, an observed change in the electronic properties of a-Si under illumination, is frequently attributed to this dynamic action of hydrogen. A typical observation of the Staebler-Wronski effect is the degradation of efficiency with time in solar cells under illumination. The illumination changes the ratio of Si-Si to Si-H bonds creating more dangling bonds and hence increasing defect sites.

The defect density can be expressed as the number of defects of a given energy per unit volume, to give the defect density of states. However, for a dangling bond defect there are three charge states: neutral, positive and negative. A neutral state corresponds to a Si dangling bond that contains only 1 electron, and hence the Si atom remains neutral. The negative state is a Si dangling bond that has captured an additional electron, hence the charge on the Si atom is negative. Following this to its conclusion, the positive state is an empty dangling bond.

To help distinguish these states the following definitions have been used throughout this thesis. \( N(E) \) represents the number of defects per unit volume that exist with an energy, \( E \). In the case of a dangling bond, the energy, \( E \), is the energy of the neutral defect. Of this density, a fraction of such states is of a given charge state and are related by the equations below.

\[
N^+(E) = f^+(E)N(E) \\
N^0(E) = f^0(E)N(E) \\
N^-(E) = f^-(E)N(E)
\]  

(4-1)
where \( N(E) \) is the density of dangling bonds, that if they remained neutral would occupy an energy, \( E \). \( N'(E) \) is the density of states of doubly occupied dangling bond states and again the energy, \( E \), is the energy of the neutral bond that forms the doubly occupied state. Finally \( N''(E) \) is the density of states of empty dangling bonds (which can be considered to contain two holes) and \( N^0(E) \) is the density of singly occupied dangling bonds. The occupancy functions \( f^0, f^- \) and \( f^+ \) give the fraction of dangling bond states that can be found in a given charge state. These functions can be derived on a first principle basis from statistical mechanics, as has been done in Appendix B.

It is important to note that throughout the energy is taken as that of the neutral dangling bond. In fact, a \( D^+ \) state has no energy associated with it and a \( D^- \) state has an energy of \( 2E + U \) associated with it, where \( U \) is a correlation energy caused by the repulsion of the electrons.

**Section 4.3 - The Defect Pool Model**

The defect pool model of a-Si attempts to thermodynamically deduce the structure of the dangling bond density. Powell and Deane\(^2\) have given a good review of the early defect pool models and have made a significant improvement to these models. They begin by considering the following sets of chemical reactions.

\[
\text{WB} \leftrightarrow (2\text{D}) \quad (4-2)
\]

\[
\text{Si-H} + \text{WB} \leftrightarrow (\text{D} + \text{Si-H}) + \text{D} \quad (4-3)
\]

\[
2\text{Si-H} + \text{WB} \leftrightarrow (\text{Si-H-H-Si}) + 2\text{D} \quad (4-4)
\]

Here WB represents a weak Si-Si bond, D represents a dangling bond and the brackets indicate species that cannot move apart (even if not chemically bonded). The
derivation of this model has been carried out in full in Appendix B-2, the following equations are the result and describe the electronic density of states for dangling bond defects.

\[
N(E_D) = \gamma \left[ \frac{2}{f^{\ominus}(E_D)} \right]^{\frac{\rho kT}{E_v}} P \left[ E_D + \frac{\rho \sigma^2}{2E_v} \right]
\] (4-5)

In the above equations, the following definitions have been used:

\[
P(E_D) = \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{(E_D - E_p)^2}{2\sigma^2} \right]
\] (4-6)

\[
\gamma = \left[ \frac{N_v 2E_v}{2E_v - k_b T} \right]^\rho \left[ \frac{i}{2H} \right] \rho^{-1} \exp \left[ -\rho \left( \frac{E_p - E_v - \rho \sigma^2}{2E_v} \right) \right]
\] (4-7)

\[
\rho = \frac{2E_v}{(2E_v + ik_b T)}
\] (4-8)

where \(E_v\) is the characteristic energy of the valance band tail, \(N_v\) is the density of states of the valance band tail at the mobility edge, \(i\) is the number of Si-H bonds participating in an interaction and \(H\) is the density of hydrogen. As is usual, \(k_b\) is Boltzmann’s constant and \(T\) is the temperature in Kelvin. \(\sigma\) is the width of the Gaussian distribution which describes the weak bonds in the material and \(E_p\) is the most probable energy for such a bond. \(E_v\) is the valance band edge energy and \(f^{\ominus}(E)\) is the occupancy function for the neutral bonds (see Appendix B, equation (B-19)). Special care has been taken to distinguish \(E_D\) as the energy of the neutral defect, which when empty or full becomes associated with a charged state. This is particularly important when considering the negative dangling bond defect energy.
This density of states can then be converted into a single electron density of states using the equations given above and in appendix B, see equations (4-5), (B-18), (B-19) and (B-20). This allows the electron density of states to be predicted from first principles. A few of these variables are not well defined, but they can be related to measured variables via simulation and so this is not a great disadvantage of the theory. In fact, the theory has proven to be very powerful in predicting effects such as reverse-bias annealing and note that the effects of hydrogen are implicit in the equations.

Figure 4-2 shows a defect distribution that has been calculated using the above formula and the defect pool parameters used to obtain this distribution are given in Table 4-1. Note however, that in the derivation of the dangling bond statistics the energy used is not that of the electron but that of the neutral defect. Consequentially, $N'(E)$ is not the density of negative dangling bonds at energy $E$. Rather, it is the density of negative dangling bonds generated by neutral defects with energy, $E$. Because the defect contains two electrons and these are correlated, these defects actually lie at an energy of $2E+U$. Such a distribution is illustrated in Figure 4-3.
Figure 4-2: The mid-gap density of states calculated using the Powell and Deane\(^2\) defect pool model and the parameters given in Table 4-1.

Figure 4-3: The mid-gap density of states calculated as with the above case but this time with the energy being the energy of charged state, not the defect.
Thus far, the dangling bond model provides an elegant and useful model of the a-Si density of states. However, to be applied to a device simulation it would be necessary to be able to derive some sort of relationship for the free carrier concentration, particularly in a non-equilibrium situation. In effect a relationship of the form given by equations (2-18) to describe the quasi-fermi level. In addition, a means of expressing the recombination function must be found.

In Appendix B, it is shown that the Fermi-Dirac distribution should still hold for the free carrier concentrations. As the dangling bond model uses the same distributions, it follows that the free carrier concentrations are given by equations (4-9a) and (4-9b).

\[
n = \frac{g(E_c)}{1 + \exp\left(\frac{E_c - E_F}{k_B T}\right)} \tag{4-9a}
\]

\[
p = \frac{g(E_v)}{1 + \exp\left(\frac{E_v - E_F}{k_B T}\right)} \tag{4-9b}
\]

where \(g(E)\) is the density of states at energy \(E\), \(E_F\) is the Fermi energy, \(E_c\) is the conduction band energy, \(E_v\) is the valence band energy, \(n\) is the electron concentration and \(p\) is the hole concentration.
The Fermi energy should be that used in calculating the dangling bond density of states, in addition the hole density of states is given, as \( N_{v_0} \), in Table 4-1. From this, the free hole concentration is calculated to be about 5000 cm\(^{-3}\) at 300 K, a low value. Now if charge neutrality is applied, then equation (4-10) should give the electron concentration.

\[
n - p + \int_{E_v}^{E_c} (D^- + D^+)dE = 0 \tag{4-10}
\]

This leads to an electron concentration of \( 1 \times 10^{17} \) cm\(^{-3}\). Using the above equations to convert this into a density of states gives a near impossible value of \( 1 \times 10^{32} \) cm\(^{-3}\). Clearly, the theory fails when steps are taken towards describing the recombination function. An examination of appendix B reveals that the derivation of the occupancy functions relies on the assumption that the dangling bond and conduction band densities can be minimised independently. Some thought suggests that this assumption is probably incorrect. The free carrier density and the dangling bond density are dependent, due to the trapping and emission of carriers. Either the system has to be described as a non-equilibrium system or else an expression must be found describing the number of ways of arranging electrons and hydrogen into the various states that can be minimised. The most intuitive solution is to consider the dangling bond recombination as a non-equilibrium system. Such approaches are discussed in the follow sections.

**Section 4.4 - The Rate Equations**

An alternative approach to the dangling bond technique is to examine the rate equations and treat the system as being in a dynamic equilibrium. This idea has a lot
of appeal as it follows the techniques used to reach the Shockley-Read-Hall model discussed in Chapter 2.

Figure 4-4 illustrates the possible interactions of a dangling bond with the charge carriers in the valence and conduction bands. It is possible to empirically describe each of these processes by a set of rate equations; these are shown in Table 4-2. The origin of these equations and the forms for the capture and emission coefficients are discussed in Appendix C.

![Diagram of recombination paths in amphoteric dangling bond states. Probabilities for the various interactions are given in Table 4-2.](image-url)
### Table 4-2: Rate equations for various types of dangling bond recombination.

<table>
<thead>
<tr>
<th>Equation Number</th>
<th>Recombination Path</th>
<th>Rate of Recombination</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4-A1)</td>
<td>( {D^- - e} \rightarrow D^0 )</td>
<td>( e_n^-(E) N^-(E) )</td>
</tr>
<tr>
<td>(4-A2)</td>
<td>( {D^- + h} \rightarrow D^0 )</td>
<td>( e_p^-(E) pN^-(E) )</td>
</tr>
<tr>
<td>(4-A3)</td>
<td>( {D^0 + h} \rightarrow D^+ )</td>
<td>( c_p^0(E) pN^0(E) )</td>
</tr>
<tr>
<td>(4-A4)</td>
<td>( {D^0 - h} \rightarrow D^- )</td>
<td>( e_p^0(E) N^0(E) )</td>
</tr>
<tr>
<td>(4-A5)</td>
<td>( {D^0 + e} \rightarrow D^- )</td>
<td>( c_n^0(E) nN^0(E) )</td>
</tr>
<tr>
<td>(4-A6)</td>
<td>( {D^0 - e} \rightarrow D^+ )</td>
<td>( e_n^0(E) N^0(E) )</td>
</tr>
<tr>
<td>(4-A7)</td>
<td>( {D^+ + e} \rightarrow D^0 )</td>
<td>( c_p^+(E) nN^+(E) )</td>
</tr>
<tr>
<td>(4-A8)</td>
<td>( {D^+ - h} \rightarrow D^0 )</td>
<td>( e_p^+(E) N^+(E) )</td>
</tr>
</tbody>
</table>

It is expected that in equilibrium the recombination rate must be equal to zero. By considering each of the reactions in Figure 4-4, the following equations can be arrived at. Note that equations (4-11), (4-12) and (4-13) all apply only to a dangling bond at a specific energy. In an amorphous solid, there is expected to be a continuous distribution of such defects. However, it is useful to approximate such a continuous distribution by a discrete distribution. For any given energy in such a distribution, there exists a set of three equations of the form given by (4-11), (4-12) and (4-13).

\[
\frac{\partial N^0(E)}{\partial t} = c_n^-(E) pN^-(E) + e_n^-(E) N^- (E) + e_p^+(E) N^+(E) + c_p^+(E) nN^+(E) - c_n^+(E) nN^+(E) - c_p^0(E) pN^0(E) - e_n^0(E) N^0 (E) - e_p^0(E) N^0 (E)
\]  

(4-11)
Dangling Bond Recombination in a-Si:H

\[
\frac{\partial N^-(E)}{\partial t} = c_n^0(E)nN^0(E) + e_p^0(E)N^0(E) - c_p^-(E)pN^-(E) - e_n^-(E)N^-(E) 
\]

\[
\frac{\partial N^+(E)}{\partial t} = c_p^0(E)pN^0(E) + e_n^0(E)N^0(E) - e_p^+(E)N^+(E) - c_n^+(E)nN^+(E) 
\]

These equations neglect any changes in the defect density due to the interaction of hydrogen. By simply considering the reactions that affect the free carrier concentrations, equations (4-14) and (4-15) can be arrived at. These describe the rate of change of the free carrier concentrations. It is important to note that the summation is in effect the recombination function for the semiconductor equations.

\[
\frac{\partial n}{\partial t} = G - \sum_{\text{all energies}} \left( c_n^+(E)nN^+(E) + c_n^0(E)nN^0(E) - e_n^-(E)N^-(E) - e_n^0(E)N^0(E) \right) 
\]

\[
\frac{\partial p}{\partial t} = G - \sum_{\text{all energies}} \left( c_p^-(E)pN^-(E) + c_p^0(E)pN^0(E) - e_p^+(E)N^+(E) - e_p^0(E)N^0(E) \right) 
\]

For equilibrium, all of these equations should be equal to zero. By noting (from equations (4-2) to (4-4)) that \( f = 1 - f^- - f^+ \) and if we assume that the dangling bond density is also in equilibrium it can be seen that:

\[
\frac{\partial}{\partial t} (f^0N_D) = \left( \frac{\partial}{\partial t} N_D - \frac{\partial}{\partial t} (f^-N_D) - \frac{\partial}{\partial t} (f^+N_D) \right) 
\]

This means that, given the above relationship between the occupancy functions, the rate equation governing the neutral dangling bonds is simply a linear superposition of the other rate equations. This allows (4-12) and (4-13) to be re-written as a pair of simultaneous equations.

\[
\alpha_1 - \alpha_1 f^+ - (\alpha_1 + \alpha_2) f^- = 0 
\]
\[ \alpha_3 - (\alpha_3 + \alpha_4)f^+ - \alpha_3f^- = 0 \] (4-18)

Here we have introduced the variables \( \alpha_1, \alpha_2, \alpha_3 \) and \( \alpha_4 \) as shorthand for the following variables, which we note, are always positive.

\[
\begin{align*}
\alpha_1 &= c_n^0 n + e_p^0 \\
\alpha_2 &= c_p^- p + e_n^- \\
\alpha_3 &= c_p^0 p + e_n^0 \\
\alpha_4 &= c_n^0 n + e_p^+ 
\end{align*}
\] (4-19)

Given that in equilibrium \( n \) and \( p \) are constant, then this is soluble for \( f^+, f^- \), and implicitly \( f^0 \), for a given set of carrier densities. From this, the following equations predict the occupancy of charge states for a given free carrier and defect density.

\[
f^+ = \frac{1}{1 + \frac{\alpha_4}{\alpha_3} \left( 1 + \frac{\alpha_1}{\alpha_2} \right)}
\] (4-20)

\[
f^- = \frac{1}{1 + \frac{\alpha_2}{\alpha_1} \left( 1 + \frac{\alpha_1}{\alpha_4} \right)}
\] (4-21)

\[
f^0 = 1 - \frac{1}{1 + \frac{\alpha_4}{\alpha_3} \left( 1 + \frac{\alpha_1}{\alpha_2} \right)} \cdot \frac{1}{1 + \frac{\alpha_2}{\alpha_1} \left( 1 + \frac{\alpha_3}{\alpha_4} \right)}
\] (4-22)

Using the equations (4-20), (4-21) and (4-22), an equilibrium occupation can be calculated for a given set of free carrier concentrations and defect densities. Note that there is no guarantee that with an arbitrary free carrier concentration the band tails are in equilibrium with the defect density. Ensuring this requires that equation (4-15) is also solved. This problem is more complex and will be returned to later.

Figure 4-5 shows a defect density calculated using equations (4-20) to (4-22) to calculate the occupancy. These have been combined with equation (4-5), and the values from Table 4-1, to give the actual density of states. The capture and emission
coefficients in appendix C have been used and the free carrier concentrations are chosen as \( n = p = 10^{10} \text{ cm}^{-3} \) (a value taken from Figure 3-7). Table 4-3 summarises the values used in this calculation. Note that whilst the tail states of the distribution take on a similar form the mid-gap density of states is significantly different. It is both broader and flatter in shape. In fact, this is due to the need to calibrate the simulations with the various parameters (i.e. cross sections) and the arbitrary choice of a free carrier concentration.

Figure 4-5: A density of states calculated using equations (4-20) to (4-22) with a free carrier density of \( 10^{10} \text{ cm}^{-3} \).

Hubin et al.\textsuperscript{3} used a similar technique to arrive at expressions like (4-20) to (4-22) and then with the aid of further approximations they arrive at equation (2-46) for the recombination term. One of Hubin et al's.\textsuperscript{3} assumption is to neglect the emission coefficients. Whilst this can be a good assumption at high free carrier concentrations, this is not the case for low bias. For convergence of the drift diffusion equations, the
boundary conditions must be ramped slowly from a zero bias solution (though not a necessary condition, i.e. convergence is still possible, just not guaranteed). Hence, this assumption is a poor one to use when attempting solution of the drift-diffusion equations.

In addition, it can be seen that eliminating the emission coefficients produces occupancy functions with no energy dependence. This contradicts the defect pool model and the occupancy functions derived in appendix B. Consequentially we retain the emission coefficients in our calculations.

It is also odd to note that the occupancy does not depend upon the density of defects. However, looking at the global problem it becomes obvious that the defect density affects n and p. This is made clear by considering charge conservation, i.e.:

$$n - p + (f^- - f^+)N_d = 0 \quad (4-23)$$

Rearranging equation (4-23) to leave only the n term on the right hand side, and then differentiating with respect to time leads to equation (4-24).

$$\frac{\partial n}{\partial t} = \frac{\partial p}{\partial t} + \left( \frac{\partial f^-}{\partial t} - \frac{\partial f^+}{\partial t} \right)N_d + (f^- - f^+)\frac{\partial N_d}{\partial t} \quad (4-24)$$

If the defect density is invariant with time then this is a superposition of the rate equations. This allows equation (4-15) to be neglected. Then using equation (4-23) to eliminate n from the rate equations, we can arrive at the following forms for the rate equations.

$$\frac{\partial f^- (E)}{\partial t} = c_n^0(p - (f^- - f^+)N_d)(1 - f^- - f^+)
+ e_p^0(1 - f^- - f^+ - c_p pf^- - e_n f^-) \quad (4-25)$$
Alas at this point analytical techniques fail because the full set of equations is coupled. Hence, there is no general algebraic solution for the defect density and the free carrier density. Numerical techniques have been tried, for example the Newton-Raphson method, however these were always found to converge to the trivial solution.

To prevent convergence of the equations on the trivial solution, it would be necessary to define some energy for the system that must be conserved. However, this is in fact what is done by the defect pool model and statistical mechanics then predicts a free carrier concentration. Hence, a vicious circle arises. Only by assuming something about the system, such as free carrier concentration or the total energy for the system, can this approach be made to work. This returns us to equations (4-20) to (4-22). However, this demonstrates that the distribution of carriers in the defect states is coupled to the free carrier concentration.

To summarise then, the rate equation approach gives a method whereby the occupancy functions can be calculated for a given defect density. However, it is quite possible, that the solutions arrived at could be out of equilibrium with the conduction and valence states. This is a reflection of the problem with the defect pool model, which also fails to predict both the free carrier and defect density equilibrium concentration.
To get around this problem it is necessary to incorporate the rate equations for the band tails. This leads to a problem that is non-convergent using standard numerical techniques. Hence, even if a robust technique was available it is likely that the computation time involved would be prohibitive for any practical simulation work.

<table>
<thead>
<tr>
<th>Capture Cross-Sections</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Electrons</strong></td>
</tr>
<tr>
<td>$D^0$</td>
</tr>
<tr>
<td>$D^+$</td>
</tr>
<tr>
<td>$D^-$</td>
</tr>
<tr>
<td>$\sigma_n^0$ (cm$^{-2}$)</td>
</tr>
<tr>
<td>1x10$^{-17}$</td>
</tr>
<tr>
<td><strong>Holes</strong></td>
</tr>
<tr>
<td>$D^0$</td>
</tr>
<tr>
<td>$D^+$</td>
</tr>
<tr>
<td>$D^-$</td>
</tr>
<tr>
<td>$\sigma_p^0$ (cm$^{-2}$)</td>
</tr>
<tr>
<td>1x10$^{-17}$</td>
</tr>
</tbody>
</table>

**Escape Frequency**, $\omega = 1x10^{12}$ sec$^{-1}$

**Thermal Velocities**, $V_n = V_p = 1x10^7$ cm/s

**Density of free electrons**, $n = 1x10^{10}$ cm$^{-3}$

**Density of holes**, $p = 1x10^{10}$ cm$^{-3}$

Table 4-3: Values used in solving the occupancy equations (4-20), (4-21) and (4-22).

Note however, that the rate equation approach is increasingly valid in cases when the band tail states are expected to be out of equilibrium with the defect states. Then the exact nature of the defect density tends to become less important. However, solutions found near equilibrium are likely to be inaccurate.
Section 4.5 - A Dynamic Dangling Bond Model

In the previous section, we have described methods that allow the calculation of the dangling bond density of an equilibrated a-Si material. However, for most applications, the question of real interest is what happens away from equilibrium. For example, if an intense light is shone upon a sample, how does the defect density change, and what does it become? In addition, if a non-equilibrium situation can be solved and then left to relax, equilibrium solutions could be found. This would provide a solution to many of the problems of the above models and allow for a universal dangling bond model.

Such an approach was investigated for the modelling work in this thesis. This section describes the outline to such a technique and explains the difficulties in its development that lead us to dismiss this technique as of little or no use.

First, it is assumed that the system is in equilibrium and that the initial defect density is known. An obvious approach is to then integrate the rate equations, (4-11) to (4-15), with respect to time. However, this falls into the trap of interpreting the rate equations as ordinary differential equations instead of partial differential equations. Actually, the equations are heavily coupled; changing the defect density at one level will affect all of the other levels. As mentioned earlier, this coupling arises because of the conversion of charged states and because of carrier exchange with the conduction bands.

It is instructive to recast the problem in general terms to carry out a mathematical analysis. We will consider an arbitrary distribution of charge represented by the function $N(E,s,t)$. Where $E$ is the energy, $s$ is the charge state and $t$ is the time. If it is
assumed that we know this function totally at an arbitrary time \( t = t_0 \) then using a Taylor series expansion we can express the function at some later time \( t = t_0 + \delta t \). This then gives us equation (4-28).

\[
N(E, s, t_0 + \delta t) = N(E, s, t_0) + \delta t \frac{\partial}{\partial t} N(E, s, t_0) + (\delta t)^2 \frac{\partial^2}{\partial t^2} N(E, s, t_0) + \ldots + (\delta t)^m \frac{\partial^m}{\partial t^m} N(E, s, t_0) + \ldots
\]  

(4-28)

where \( m \) is an index that should run from 0 to infinity in the full expansion. This allows us to use the partial differential equations (4-11) to (4-15) to find the solution at any arbitrary time interval with one caveat - the series must converge! Clearly as \( \delta t \) tends to 0 then, the series will probably converge. But this is not very helpful as if the series can only be used for infinitesimally small steps, then we will not be able to solve a reasonable number of points for numerical simulation of devices. Unfortunately, any mathematical analysis of the convergence of this series is complicated by the fact that the partial derivatives are coupled to all of the other equations and we rapidly end up with a morass of simultaneous equations.

We can at least take a pragmatic approach to this problem. We can solve for an arbitrary number of differentials and then observe if there are signs of divergence. If not, then the time step can be reduced in size until the Taylor series is accurate within our requirements. Clearly this does not preclude the possibility of higher derivatives 'blowing' up to give a rapidly diverging series, but the complexity of the equations leaves us with little choice.

Table 4-4 shows the arbitrary \( i^{th} \) derivative. In calculating this derivative, \( i = 1 \) is taken to be the rate equation or the first derivative of charge with respect to time, \( i = 2 \)
corresponds to the second derivative and so forth. This means that the binomial series can be used to calculate the \( i \)th derivative from the previous derivatives. Here we are using the convention that the zeroth derivative of a function is in fact the function itself.

<table>
<thead>
<tr>
<th>Equation Number</th>
<th>Recombination Path</th>
<th>The ( i )th derivative of charge.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4-B1)</td>
<td>( {D^--e} \rightarrow D^0 )</td>
<td>( e_n^-(E) \frac{\partial^{i-1}}{\partial t^{i-1}} [N^-(E)] )</td>
</tr>
<tr>
<td>(4-B2)</td>
<td>( {D^0 + h} \rightarrow D^0 )</td>
<td>( e_n^0(E) \frac{\partial^{i-1}}{\partial t^{i-1}} [N^0(E)] )</td>
</tr>
<tr>
<td>(4-B3)</td>
<td>( {D^0 + h} \rightarrow D^+ )</td>
<td>( e_n^0(E) \frac{\partial^{i-1}}{\partial t^{i-1}} [N^0(E)] )</td>
</tr>
<tr>
<td>(4-B4)</td>
<td>( {D^0 - h} \rightarrow D^- )</td>
<td>( e_n^0(E) \frac{\partial^{i-1}}{\partial t^{i-1}} [N^0(E)] )</td>
</tr>
<tr>
<td>(4-B5)</td>
<td>( {D^0 + e} \rightarrow D^- )</td>
<td>( e_n^0(E) \frac{\partial^{i-1}}{\partial t^{i-1}} [N^0(E)] )</td>
</tr>
<tr>
<td>(4-B6)</td>
<td>( {D^0 - e} \rightarrow D^+ )</td>
<td>( e_n^0(R) \frac{\partial^{i-1}}{\partial t^{i-1}} [N^0(E)] )</td>
</tr>
<tr>
<td>(4-B7)</td>
<td>( {D^+ + e} \rightarrow D^0 )</td>
<td>( c_n^+ \left{ \frac{(i-1)!}{((i-1)-j)!j!} \left[ \frac{\partial^j}{\partial t^j} [N^+] \frac{\partial^{(i-1)-j}}{\partial t^{(i-1)-j}} [N^+(E)] \right] \right} )</td>
</tr>
<tr>
<td>(4-B8)</td>
<td>( {D^+ - h} \rightarrow D^0 )</td>
<td>( e_n^+(E) \frac{\partial^{i-1}}{\partial t^{i-1}} [N^+(E)] )</td>
</tr>
</tbody>
</table>

Table 4-4: Arbitrary derivatives using the binomial expansion to calculate the capture rates. These equations can be used in conjunction with equation (4-21) to calculate the time dependence of dangling bond occupation. Note that the binomial expansions only exist for \( i=1 \) or greater.

From the equations (4-B1) to (4-B8), plus the Taylor series expansion of the charge density (equation (4-28)), it is possible to calculate the defect distribution after an arbitrary time, even under non-equilibrium conditions. This assumes that the evolution with time is smooth (i.e. the derivatives exist and are finite) and that the Taylor series will converge.
However, it can be seen from the form of the above equations that the equations will probably diverge. The conditions for convergence are that the rate of change of defect density is less than the emission coefficients. Also, that the product of the free carrier concentration and the rate of change of defect density is less than the capture coefficients. That this was so was confirmed by calculating equations (4-B1) to (4-B8) for the defect density given in Figure 4-2. The results gave second order derivatives of the order of $1 \times 10^{30}$ carriers/sec$^2$. This would require a time step less than $1 \times 10^{-15}$ sec to prevent the series from diverging. Clearly, such a time step is too small to derive useful results. In addition, the large differences between numbers makes the method very numerically unstable.

Section 4.6 - Conclusions

We have examined the predominant models for dangling bond defects in a-Si with the aim to implementing a fuller recombination model in the device simulations. The dangling bond model fails to give realistic values for the free carrier concentration. Thus, it cannot be used as part of a drift-diffusion calculation.

The rate equation approach provides a better system for calculating a recombination function. However, like the dangling bond model there appears to be no self-consistency between the calculated defect density and the free carrier concentration. Hence, we conclude that near equilibrium the method cannot predict dangling bond densities. However, given a consistent set of dangling bond densities and free carrier concentrations is can at least be used to arrive at a recombination function.

Attempts to find a consistent solution to the rate equation problem by considering the transient behaviour have failed. This is because the resulting sets of coupled partial
differential equations are divergent except under very strict convergence conditions that correspond to points close to equilibrium.

**Section 4.7 - References**


Chapter 5 Peripheral Design Concerns and Conclusions

Section 5.1 - Introduction

The principle aim of this thesis has been to examine new photodiode structures with a view to their incorporation into an a-Si:H X-ray detectors. Consequentially, that has been the prime concern of the previous chapters. However, there are some interesting points to be made about the design of the other detector components, and work that could be carried out in the future. This chapter begins by examining these other components and the possibilities for future work in this area. The chapter then ends with a summary of the conclusions of the work carried out on a-Si X-ray detectors.

Section 5.2 - Peripheral Components

5.2.1 Introduction

The work that has been carried out thus far has concentrated on the a-Si photodiode as being the key component in the medical X-ray detector. However, there are five identifiable key components in addition to the photodiode each of which has their own set of issues that need to be addressed in the design of the detector. These are as follows:

- The Conversion Layer
- The Pixel Read Out Electronics
- The Digital Switching Electronics
- The Digital Readout to a Computer System
- The Calibration of data from the array
Each of these points is discussed in the following sections, in the order given above. The influence of these components is assessed and the direction of any future work on these components is suggested.

5.2.2 The Conversion Layer

When an electromagnetic wave passes through a material, the electric field can couple to the charge of the atomic electrons, thus ionising the atoms. An ideal conversion layer would completely absorb all of the X-ray's energy in this manner resulting in a 100% conversion of the X-ray's energy into electronic signal. However, some of the ionised atoms will recombine with the free electrons causing a loss of signal. In addition, these recombining species constitute an electric dipole capable of radiating more X-rays. These fluorescent X-rays may escape the conversion layer or be re-absorbed to generate electrons. Since a fluorescent X-ray can be emitted in any direction, this means that the electronic signal generated may be a considerable distance from the initial photon interaction site.

This conversion of photons to electrons and back again is a chain of events known as the electro-magnetic shower. Its importance to the operation of the conversion layer is twofold: Firstly, the possibility exists that the secondary X-rays and electrons generated in the shower may escape the conversion layer. This causes a loss of signal and the escaping radiation may damage or interfere with the electronics surrounding the conversion layer. Secondly, the final signal generated by the X-ray's interactions may be a considerable distance from the original interaction site of the X-ray. This problem is exacerbated by the long mean free path between photon interactions and the fact that a single X-ray may undergo several interactions before finally losing all
Peripheral Design Concerns and Conclusions

of its energy. This second point contributes to a loss of spatial resolution in the
conversion layer.

It is possible to model this process using Monte Carlo techniques. This technique
begins by generating a particle and a set of probability distributions are randomly
'sampled'* to determine what set of interactions the particle would undergo. The
interaction process is repeated for each particle until the particle is annihilated, it
leaves any region of interest or it loses sufficient energy to be ignored. This
procedure can be repeated for other particles until a statistically significant result is
obtained.

The advantages of Monte Carlo techniques are that they directly model the underlying
physics of a system and that they can also reveal statistical variations (i.e. noise) and
so are easily compared with experimental results. However, such simulations
frequently require several thousand particles to begin to achieve a statistically
significantly result: hence, the simulations can be very slow. Another problem is that
the probability distributions required for the simulations can be hard to measure
directly. Finally, the simulations are quite complex and it is hard to intuitively see
how changing a set of parameters (such as particle energy) will impact on the result.

There are several pre-written codes available for Monte Carlo study of radiation
transport, such as EGS¹ and MCNP². These codes are complex and are designed
primarily for high energy physics. At the energy regime of diagnostic X-rays, the

* Sampling has a specific meaning in Monte Carlo techniques. The principle aim of
'sampling' is to find a way of using a flat, normalised random number to produce a
sequence of random numbers that can match a required probability distribution.
Peripheral Design Concerns and Conclusions

Simulations begin to touch the lower limits of their applicability. Consequentially, such simulations have fallen outside of the scope of this thesis.

If such simulations were to be carried out, they would answer some of the pertinent questions regarding the ultimate spatial resolution of the detector but would not tell the whole story. Optical photons generated in the scintillator would propagate randomly, causing a further spread of signal. Again, there exist pre-written codes (e.g. MCML$^3$) capable of dealing with such problems. However, the many unanswered questions regarding the nature of columnar CsI make such calculations difficult. Any work to be carried out in this direction must also be backed up by experimental work.

The aim of such Monte Carlo work would be targeted at selecting an optimum thickness for a conversion layer and to investigate the improvements in spatial resolution achieved by using a columnar CsI scintillator. The optimum thickness is a trade off. A thick scintillator maximises absorption of the incident X-rays (and prevents energetic electrons from the resulting interactions from damaging the underlying photodiode). However, this means that the resulting optical signal will diverge more and spatial resolution will be lost compared to a thin scintillator.

This trade off can be illustrated with a simple two dimensional model of light propagation. First, it is assumed that there is no absorption in the phosphor and that diffraction plays no significant role. Then it is assumed that the radiation generates light from a point source at the surface of the phosphor with an intensity $I_0$. For a plane piece of phosphor of thickness $H$, the light incident of the far surface can then be calculated using the inverse square law and the response is the given by equation (5-1) below.
Peripheral Design Concerns and Conclusions

\[ I(x) = \frac{I_0}{k(x^2 + H^2)} \]  \hspace{5cm} (5-1)

Here \( k \) is a constant and \( x \) is taken as being the lateral displacement from the point of interaction of the beam.

Figure 5-1: Illustration of how a columnar structure can help contain light within the region in which the interaction occurs. Here a column is considered to be 2\( R \) width and \( H \) thick. The interaction, i.e. the point where light is emitted, is assumed to be at a single point at \( x=0, y=H \).

Figure 5-1 illustrates a columnar structure, where the columns lie at \( x=\pm R \). A simple model of this can be constructed if it is assumed that the walls of the column are ideal diffusers, which re-radiate any incident radiation equally in all directions. Hence, any light incident on the walls of the structure is re-emitted as if from a point source. Therefore, the resulting signal along the base of the phosphor is given by equation (5-2) if secondary reflections are neglected.
Peripheral Design Concerns and Conclusions

\[ I(x) = \frac{I_0}{k(x^2 + H^2)} + I_R(x, y) \quad \text{for} \quad |x| < R \quad (5-2) \]

Here \( I_R(x, y) \) is the intensity of light due to diffuse reflections from the columns. This function can be assessed by calculating the incident flux on an infinitesimal part of the column wall and then by applying equation (5-1) to find the contribution of each such point source to the final intensity. This can then be integrated to give the total intensity at a point. Noting that there are two contributions from beams either side of the column the algebra results in equation (5-3).

\[
I(x) = \frac{I_0}{k(x^2 + H^2)} + \frac{I_0 D}{k} \int_0^H \frac{1}{ \left( \frac{R}{y} + \frac{(H-y)^2}{y^2 + (R-x)^2} \right) \left( \frac{R^2}{y^2 + (R-x)^2} \right) } dy \quad (5-3)
\]

Here \( D \) has been introduced as a number that accounts for the fraction of light lost during the scattering process. For an ideal diffuser this would be 1, typically it would be less than 0.1. In addition, for a more realistic case it should also be a function of angle of incidence of the light. However, this is neglected for simplicity.

As the calculation assumes that secondary reflections are negligible, the only contribution to light outside the column is that of the diffuser on that side of the structure. Hence, for \( x > R \) the intensity is given by equation (5-4).

\[
I(x) = \frac{I_0 D}{k} \int_0^H \frac{1}{ \left( \frac{R}{y} + \frac{(H-y)^2}{y^2 + (R-x)^2} \right) \left( \frac{R^2}{y^2 + (R-x)^2} \right) } dy \quad (5-4)
\]

Though not trivial, the integral can be carried out to give the result for the intensity of light in the region of the column where the interaction occurs. This has been done
Peripheral Design Concerns and Conclusions

using maple to produce the plots in Figure 5-2 and Figure 5-3. This data has been calculated using a thickness (H) of 50 and a column width (2R) of 200 (these values have been deliberately left unitless for ease of use). For simplicity, k and I are taken as 1 and the diffusivity constant is taken as 0.04. It should be noted that because of the assumption of an infinitesimal diffuser at the column edges, the calculation would fail approaching the edges due to the discontinuity in the functions. The first figure shows how the intensity at the base varies within the boundaries of the column and represents signal. In the second figure, what happens outside of the column is illustrated and represents crosstalk.

Figure 5-2: A plot of the intensity of light emerging from the base of a phosphor. X=+/100 marks the edges of the columnar diffuser. This plot shows both the columnar case (with R=200, H=50, k=1 and D=0.04) and a plot assuming only equation (5-1). Only one line is visible because the match is almost perfect.
Figure 5-3: A plot of the intensity at the base of the phosphor. The source lies at \( x=0 \) and a diffuser lies at \( x=100 \). This plot shows both the columnar case (with \( R=200, H=50, k=1 \) and \( D=0.04 \)) and a plot assuming only equation (5-1). Only one line is visible because the match is almost perfect. The y scale has a major unit of \( 1 \times 10^{-5} \).

The plots show little difference between the two models indicating that for \( R \gg H \) that the diffuser makes little difference. However, using this model, the thickness of the scintillator can changed, Figure 5-4 and Figure 5-5 illustrate what happens when the thickness is increased from 50 to 200. From Figure 5-4, it can be seen that the presence of the diffuser increases the collection efficiency of the scintillator over the diode compared with that with no diffuser. From Figure 5-5, it can be seen that the system with no diffuser in place shows an increased spread in the light signal into the adjacent region. The plots show clearly that with the diffuser in place the confinement of the signal is far greater than that without a diffuser in place. Obviously though this improvement has come about with the loss of signal.
Figure 5-4: A plot of the intensity of light emerging from the base of a phosphor. $X=\pm 100$ marks the edges of the columnar diffuser. This plot shows both the columnar case (with $R=200$, $H=50$, $k=1$ and $D=0.04$) and a plot assuming only equation (5-1). The upper line represents the columnar case, a greater fraction of light is collected, especially near the diffusing edges.

The results from this model have deliberately been left unitless to prevent false conclusions being drawn from the data. The model is only a simple one and multiple reflections mean that more light would leak through the diffuser than has been accounted for in this model. In addition, the increasing response at the edges of the columns would not exist in a fuller model because of the non-ideal nature of the scattering at the interface, which would have some directional dependence.
Figure 5-5: A plot of the intensity at the base of the phosphor. The source lies at x=0 and a diffuser lies at x=100. This plot shows both the columnar case (with R=200, H=50, \( k=1 \) and \( D=0.04 \)) and a plot assuming only equation (5-1). Clearly there is virtually no leakage when a diffuser is in place. In this plot the y scale has a major unit of \( 2 \times 10^{-6} \).

5.2.3 The Pixel Readout

The simulations in the thesis have focused on the operation of the a-Si photodiode. However, it is also possible, using the SILVACO software, to simulate the operation of an a-Si TFT. This can then be incorporated in an equivalent circuit, such as that shown in Figure 2-14 and used as part of a SPICE simulation. SPICE (Simulation Program with Integrate Circuit Emphasis) is an industry standard simulator, developed at Berkeley, for simulating non-linear circuit elements. Such a simulation could also be used to determine the effectiveness of the integrating capacitor design and to evaluate the various sources of electronic noise in the circuit.
Peripheral Design Concerns and Conclusions

Also of interest is the readout time, the noise introduced by the pixel and the introduction of stray capacitances from TFT and data lines.

A feel for the figures involved, and the problems that exist can however be reached without recourse to such simulations. We can take typical values of the TFT resistance to be $10^{15} \, \Omega$ in the off state and $10^6 \, \Omega$ in the on state. An estimate of the capacitance of the data line can be obtained if it is assumed that along a given data line there are 1000 TFTs (i.e. a 10 cm array), each with a capacitance of 5 fF. An additional 5fF is generated at each pixel by the gate line/ drain line overlap capacitance. The resulting RC time constant is of the order of 20 $\mu$s. To ensure the readout of approximately 99% of the signal the data lines must be left open for 5 times this time constant. The result is a row readout time of 100 $\mu$s. For 1000 pixels this would give a frame rate of 10 Hz. This is too low for video rate imaging but sufficient for many medical purposes. However to improve this to video rate, the number of pixels would have to be reduced or the 5 time the time constant condition relaxed. Readout rate could be improved by multiplexing the readout system to read out several rows at a time.

The minimum signal on the photodiode can be taken to be typically of the order of 1 fC (from a 1 second integration time over a 1 fA dark current). However, now note that the gate and the diode are capacitively coupled through the gate/source overlap, similarly for the data lines. Again, this capacitance is estimated to be approximately 5 fF. Hence, when turned off there is a drain to gate voltage of $\sim -5V$ and so there would be 25 fC of charge stored at this junction. This charge would normally drift onto the larger capacitance of the photodiode and/or storage capacitor. When turned
on the gate voltage will rise to $\sim +5V$ with respect to the source injecting 50 fC onto the diode/gate capacitor. The result of this is that the individual pixel will be subject to a pedestal charge of $\sim 50$ fC due to the overlap capacitances of the TFT. This emphasises the need for a small operating range on the gate voltage of the TFT to reduce this pedestal charge and consequent 'reset' noise.

In addition to the pixel reset noise, there are several other sources of noise that should be considered:

- **Shot Noise in the photodiode** - there are shot noise contributions ($2qIdf$) due to the signal current, the photodiode leakage current, the TFT leakage current and the 'readout' current. Generally, these noise sources do not dominate.

- **Thermal Noise** - this arises due to thermal fluctuations in the resistive parts of the system. A significant contribution to the total system noise arises from the TFT's resistance when 'on'. Typically the TFT will have a thermal noise current of $\sim 2kTg_D + k/f$ where $g_D$ is the transconductance of the TFT and $f$ is the frequency. A typical figure$^4$ for this noise current is $5 \times 10^{-8} + 5 \times 10^{-13} /f$ A$^2$/Hz.

- **Input Voltage Noise from the Readout Preamplifier** - because the readout electronics are not perfect, the input of the preamplifier will create some electronic noise. This results in a loss of signal-to-noise ratio and is usually the dominant noise source in practical imaging arrays.

A typical noise level for a-Si:H active matrix imaging arrays is around 1000 electrons/pixel with this noise resulting from TFT thermal noise and preamplifier input noise. It is very unlikely that the photodiode noise will dominate unless it has an exceptional leakage current.
5.2.4 Digital Switching Electronics

At first sight, the electronics for switching the TFT array are of little interest. This is because the gates of the TFTs should be insulating and hence any signal along the gate lines will not contribute directly to the pixel output. However, returning to the point of capacitive coupling of the gate lines to the pixels, if the gate driver is noisy, then signal can be induced onto the pixel. 10mV of noise will translate into 0.05fC of charge injected into the diodes. In addition, if the gates are overdriven when switched on the amount of charge induced on the junctions increases in proportion. Consequentially, the switching electronics must be designed to give a smooth, low noise pulse whose range of operation is always the same and limited to the minimum required for operation.

5.2.5 Computer Interface and Data Readout

If it assumed that each pixel has a dynamic range of $\sim 10^4$ then to carry that information would require 16 bits of digital information. Each such translated bit of data would then have to be transported to an external system. At this point, it is no longer possible to operate parallel signal lines. This means that for an array of 1000 x 1000 pixels a phenomenal 7 MBytes/s of data must be transported to a remote site. This is a difficult task requiring a fibre optic cable to carry the data and an advanced multiplexing system to convert the data into a serial stream of data and store it in memory. When compared to seek times of 10 ms on computer hard drives it can be seen that this rate of data transport requires large amounts of fast memory and plenty of storage space. Realistically the data should be carried directly into memory and then stored permanently by dedicated systems.
5.2.6 Pixel Calibration

In addition to the difficulties associated with getting data from a digital X-ray detector there are issues regarding the interpretation of this data. Firstly, the signal received from each pixel will be different for the same incident signal. The primary reason for this is process variation across the active matrix. Due to variations in thicknesses of materials, alignment tolerances and material quality, each pixel will have a different offset and possibly a different response. For example, the threshold voltage across the array will probably vary by a few millivolts.

Another source of non-linearities is the distributed resistances of the data and array lines. Diodes near the centre of the array would lose more signal through the resistance of the data lines than those at the edges of the array. Hence, another reason for individual pixels to have different responses.

All of these effects mean that an effective way to calibrate the array must be found. The obvious technique is to measure a background image, which can then be subtracted from future images. This will not however account for non-linearities in the pixel’s response, but only for offsets. To cope with non-linearities would require several images to be acquired under known illumination conditions. Then some form of interpolation could be used to calibrate the individual pixels. However, merely subtracting a 16 bit number from 7 MBytes of data with a 300 MHz Pentium could be expected to take over 20 ms. This massive feat of real time processing is not beyond the capabilities of Digital Signal Processors, which could be applied to this task.

However, the problem of calibration is a much more difficult one. Not only can the signal be a function of dose, but there are also several other sources of instability. For example, variations in temperature will change the currents in the semiconductor and
Peripheral Design Concerns and Conclusions

hence the response of pixels. Radiation damage and illumination of the photodiodes will lead to further changes in the system with time. These changes then have further impacts by causing variations in the readout rates of the electronics and the image lag of the sensor. Then there are calibration issues relating to medical imaging issues, such as accounting for beam hardening.

To account for such effects, requires various models and careful thought. The main difficulty being that any model must be simple to implement mathematically if the processing is to be carried out in real time.

Section 5.3 - Electro-Optic Probe Station

5.3.1 Introduction

It would be desirable to verify the results of the simulation in this thesis with experimental data. This has not proven to be possible due to manufacturing problems with the a-Si photodiodes. However, the equipment has been assembled to carry out relevant measurements and its operation has been demonstrated. Here the design of the Electro-Optic probe station (shown in Figure 5-6) is discussed. This equipment is designed to allow a large range of measurements to be carried out to characterise both the DC and transient electrical characteristics of photodetectors. In particular the spatial variation of such properties within the photodiode. First, the hardware that has been assembled is described. This is followed by a description of the software written to operate this probe station. Remote operation is important as it allows the equipment to be optically isolated to reduce noise and it allows the use of high powered laser sources. Finally, a few results obtained using this equipment are discussed.
5.3.2 Outline of the Electro-Optic Probe Station Hardware

Figure 5-6 shows a photograph of part of the Electro-Optic probe station designed to verify the operation of focussed photodiodes and to carry out general optical and electrical measurements. A schematic is given in Figure 5-7. The aim of this equipment is to allow electrical probing of a sample whilst also applying a small
optical beam to part of the sample. The dashed-dotted lines represent the path of beams through the station. The various inputs and outputs to the station are labelled and can be matched with the equipment in Figure 5-8.

For stability, the whole probe station is mounted on an optical bench, which can be pneumatically raised to improve vibration isolation. At the centre of the probe station, lies the sample; this can be moved manually using a set of xy micropositioners directly beneath it. These micropositioners allow for adjustment of the sample with respect to the beam position. The sample can then be probed electrically using one of 4 probes. Each probe is attached to a set of xyz micropositioners and can be connected via coaxial cables to external measurement and power units. The probes are mounted on magnetic mounting blocks, these allow the probes to be moved manually, or to be remove entirely. The micropositioners allow the probes to be adjusted to within a few microns, thus allowing contact to be made on small features.

Above the probe station there is a microscope to image the sample (connected to a monitor and a computer). This allows the probes to be imaged so that they can be accurately placed on contacts. It also means that the sample can be imaged whilst under illumination and so the beam position with respect to the sample can be ascertained. The CCD means that the sample can be imaged remotely, especially for cases where high intensity laser light is used that would otherwise be dangerous to the user.

The probes and the samples can all be moved via the two XY micropositioners beneath the probe station, these are controlled by a computer. These move both the probes and the sample with respect to a stationery beam. Above the probe station is a
Peripheral Design Concerns and Conclusions

static lens which will focus light onto the sample. This lens has an adjustable height so that the focus can be adjusted for samples of various thicknesses. The final beam size is approximately 5 microns depending on wavelength and the incident light source. By shifting the xy position of the probe station the optical beam can be raster scanned across the surface of the photodiodes.

Mounts for optical detectors are placed at various points to allow for the monitoring of the beam power, transmission through the sample and optical triggering of electronic measurements.
Figure 5-7: A schematic diagram of the electro-optic probe designed and constructed for profiling the spatial/optical response of photodiode structures. The use of the various components is explained in the text and the various outputs and inputs marked refer to Figure 5-8.
This probe station can then be connected up to various measurement devices and control apparatus, which in turn are controlled by a PC. A schematic of the connections available is shown in Figure 5-8. At the centre of this diagram is the probe station, as we have described. There are three inputs available, the first is the light source. This can be a monochromatic laser beam, a broad beam optical source or a pulsed laser beam. If a broad beam source is used then this is passed through a monochromator and the wavelength of light being used can then be selected. The next two inputs are the x and y position of the sample with respect to the incident beam. The electrical probes can be either pulsed, using the pulse generator or driven from a DC source using the parameter analyser. The parameter analyser can then also monitor the current or voltage at a given probe. It can do this in both DC and transient modes. The oscilloscope can monitor the transient response of a probe. All of these devices can be monitored and controlled by the PC using the GPIB (General Purpose Interface Bus).
Figure 5-8: A schematic of the control system on the electro-optic probe station. The outputs and inputs refer to the points marked on Figure 5-7. Single solid lines represent electrical connections, dashed/dotted lines represent optical inputs from arbitrary optical sources, dashed lines represent control lines to the PC and the thick solid line indicates the GPIB interface.

Three of the four outputs can be passed on to the parameter analyser or the oscilloscope. These are the photodetectors; two of these will output a DC signal which can be monitored by either the parameter analyser or the scope. In addition, the third photodiode output is a high speed photodiode. This can be monitored on the scope or it can be used as a trigger for the scope, parameter analyser and function generator to start operation. The final output is the microscope image of the device, which can be frame grabbed to the computer if of interest. Figure 5-9 shows a photograph of some of the equipment that has been assembled for such work.
Figure 5-9: Photograph of the measurement apparatus available for use with the electro-optic probe station. With the exception of the power supplies, all of the equipment can be controlled via a remote PC.

5.3.3 Software for Controlling the Electro-Optic Probe Station

There are several reasons for wanting to automate such a system to as large a degree as possible. Firstly, if the system is automated then it can be placed in a dark room to prevent optical interference during measurements. This also has the desirable effect of reducing any danger due to stray beams. Automation also helps reduce the
possibility of static discharge disrupting the devices. Finally, many of the profiling measurements that would be made on such a system are laborious and time consuming. The data in Figure 5-11 took over 2 hours to obtain with a computer. If the experiment had used several wavelengths of light, or also scanned the y dimension then this would rapidly increase to several days.

Whilst it would have been relatively simple to write software to carry out a specific task it was felt that the software should be able to respond to the needs of several users. Hence, it should be able to offer the user all of the options that the user would have whilst working at the control panel of a piece of apparatus. However, it should also offer the user the advantage of being able to run several pieces of equipment from a single point and have the apparatus work together.

The software was written in Visual Basic 3, with some C code to interface to the stepper motor controllers. Figure 5-10 shows a screen shot of the software in operation. The menu system allows access to set ups for the various devices. For example it can be used to define the GPIB port numbers for various devices, shift the position of the stepper motor controllers or defined a waveform for use on the pulse generator. In addition, the menus allow the set ups and any scripts written to operate devices to be saved for future use. Devices can be controlled by using a simple scripting language, the script for the current experiment is displayed on the left of the screen where it can be edited. During an experiment the output screen notes what command is being run and reports any errors. Below the start and finish time for an experiment are monitored. This allows the duration of experiments to be assessed, and to check if the system has crashed during execution.
Figure 5-10: Screen shot of the integrated controller software for controlling the electro-optic probe station. Menus allow the various device interfaces to be set up and monitored. These also allow the setting of defaults for the devices. The enabled drivers check box list (top right) allow various devices to be enabled/disabled. The command box to the right allows scripts to be written to control the system.

The scripting language has been kept simple, however, given the number of different items available a full description of the language is beyond this thesis. However to illustrate the flexibility of the controller software it is worth discussing the available features. The first command of any line consists of a 1 or two letter command. The first letter defines the device of interest, the second refers to some aspect of the command or could be redundant. The selection of commands are as follows (items in italics represent user input):-
Peripheral Design Concerns and Conclusions

1. File commands. These can be used to save data, such as the position of a stepper motor or a trace from an oscilloscope channel.
   - FO \{Filenum\}rber, "Filename" to open files.
   - FC \{FileNumber\} to close files.
   - FW \{File Number\}, \{(DataList)\} to write data to files.

2. Loop Commands that allow repetition of command sequences. The loop index can also be included in file names.
   - L \{Number of Loops\} marks the beginning of a loop.
   - EL marks the end of a loop.

3. Device commands these are prefixed as below, details for individual devices are not listed as these are far to numerous.
   - H for HP4146B parameter analyser controls (e.g. start sweep).
   - O for Oscilloscope controls (e.g. set trigger).
   - G for Function Generator controls (e.g. program function frequency).
   - M for the monochromator (e.g. set wavelength to 600 nm).
   - S for the stepper motor drivers (e.g. move probe station by 10 µm).
   - A for the ADC card (e.g. sample 100 voltages).

4. Visualisation and data commands. These are V to bring up a graph to visualise the data and I to produce an image from the CCD camera.

Each of the device commands then has a selection of options that allows the users access to the operation and set up of the device. The sample script shown in Figure 5-10 operates the parameter analyser; it runs a voltage sweep from -1V to 1V with a compliance of 2 pA.
An example of the operation of this equipment would be to bias the photodiode using the parameter analyser. Then a pulsed laser beam could be brought into the device. This would activate the high speed photodiode which could be used to trigger the scope and the function generator. This would allow transient current measurements to be taken.

5.3.4 Some Results Obtained from the Electro-Optic Probe Station

A simple experiment, ran with the equipment and software, was to use a fixed wavelength DC beam to profile the optical response of a photodiode. This is the experimental equivalent of the simulations run in Chapter 3. Figure 5-11 shows the results of a test experiment used to check the equipment’s performance. Here a large area Si photodiode was placed in the probe station and reverse biased at 1 V and illuminated with a 1 mW/cm² laser beam with a wavelength of 635 nm. The device was moved in the x direction at intervals of 0.02 cm. The device was then held at a constant voltage for 1 minute to allow time for the system to stabilise. The resulting photocurrent was then measured and has been plotted as a function of position in Figure 5-11.
Figure 5-11: A plot of the current developed by a large area c-Si photodiode versus the position of the illumination creating the current. Notice how the current falls off at the edges of the photodiode.

Figure 5-12 shows some more data taken using the electro-optic probe station. Here a pulsed beam of light (635 nm laser light with a pulse width of 1 μs and a spot size of 10 μm) is incident on a GaAs detector. The detector consists of a p⁺/Intrinsic/Ohmic Contact sandwich. The detector is cleaved and biased and the laser beam used to simulate an interaction within the device. The beam is scanned across the cleaved surface of the detector and the pulse height is measured using an oscilloscope. The resulting plot of pulse height versus beam position gives a direct measure of the internal charge collection efficiency within the device.
Figure 5-12: A plot of the pulse height, from a GaAs detector, versus depth of the interaction. The interactions are simulated using a red laser beam as described in the text. The scale is set so x=0 is the p contact and x=150 is the ohmic contact. (After Paul Sellin et. al.5)

Section 5.4 - Conclusions

To conclude, we have discussed the operation and physics of a-Si photodiodes, with a view to using them as a component in a medical X-ray detector. The ability to make high quantum efficiency photodiodes in large area arrays means that this technology is suitable for this application. However, there are issues regarding charge collection from the diodes and the fill factor of detectors manufactured using this technology. In an attempt to explore these issues, a simulation model was developed for a-Si photodiodes. This model is found to replicate the principle properties of a-Si photodiodes. Namely thickness independence of diode properties, ideality factors that lie between 1 and 2 and recombination dominated reverse bias leakage.
Peripheral Design Concerns and Conclusions

This model has then been applied to a variety of diode structures in an attempt to find a technique for minimising the signal lost due to poor charge collection in low field regions of the device. It was illustrated that n-i-p diodes generally have a better performance than p-i-n ones.

Various structures were examined in an attempt to introduce a lateral field that would increase charge collection in the regions of the device not covered by the contacts. However, it was found that attempts to introduce a lateral drift field were prevented by the large signal electrode to focus electrode distance compared to the thickness of the device. Five structures were examined. The first was the junction photodiode, which relies on an additional junction to introduce a lateral field. It was found that this device could not be biased to induce a lateral field without biasing the n-i-n or p-i-p diode into reach through. The resulting current would then degrade the diode properties. However, current could be collected from the n-i-n photodiode but this was considered to be of no use due to non-linearities and difficulties in reading out from these electrodes.

The insulated gate photodiode was also seen to have no appreciable effect. It was shown that the presence of the additional contact allows collection of charge in the regions over the focus contact. However, only as far as collection could be prevented, no increase in signal was found.

The skewed photodiode was designed to induce a lateral field by offsetting the contacts. This was initially very promising, showing signs of improved collection at the edges of the photodiode. However, further analysis indicates that this signal is due to a parasitic JFET, which is being switched on by the incident light. Though this does not make the signal any less useful, such a signal would be non-linear with
respect to the intensity and wavelength of the illuminating light. The introduction of insulating contacts was found to allow some control over this effect, however turning off the parasitic JFET only results in a lower signal.

The conclusion from this was that the advanced structures would not give a significantly improved signal, given the extra processing steps involved in their production.

It was noted that the standard model used for recombination in a-Si was probably insufficiently accurate. In particular it would fail to predict the recombination function fully, this has a direct impact on the calculation of dark currents. The current techniques probably underestimate the reverse bias current. In addition, better models would be required to model the transient behaviour of the a-Si diodes, such as image lag. Fuller models would also be important in modelling temperature effects, damage from radiation and metastability. An examination of two models, the defect pool model and the rate equation approach, revealed that both models fail when applied to the problem of solving the semiconductor equations. However, a new approach to deriving the occupancy statistics from the rate equations was found.

A new approach, via the analysis of the transient problem, was examined but this was found to be numerically unstable and inapplicable.

With this we conclude that a-Si is well suited for medical X-ray applications. However, there appears to be little scope for improvement of the fill factor of these sensors. In addition, the models currently available for modelling recombination in a-Si fall short of what is required for use in simulating medical X-ray detectors. The direction that further work in this direction should take is not obvious. However, it
would seem that a reformulation of the defect pool model with a broader development of the statistics for dangling bonds may prove fruitful.

Section 5.5 - References

1 EGS4 (Electron Gamma Shower Version 4) has a long history originating from work at the Stanford Linear Accelerator. Currently A. Bielajew at the Institute for National Measurement Standards leads the efforts to maintain the code. The source code is available from the EGS4 website: http://eesun.lbl.gov/egs/egs.html, along with several other useful documents on the topic.

2 MCNP (Monte Carlo code for Neutrons and Protons) is available from the Nuclear Energy Agency, Paris, France. Further details can be found at the MCNP website: http://www-xdiv.lanl.gov/XCI/PROJECTS/MCNP/.

3 MCML (Monte Carlo Modelling of Light Transport) code was written by Lihong Wang and Steven L. Jacques and is freely available, along with manuals from the MCML website: http://ee.ogi.edu/omlc/science/software/mc.


Appendix A  An Introduction to Numerical Methods

Introduction

There is a large body of work regarding numerical techniques applied to physical problems. The greater part of this work concentrates on proving the techniques and demonstrating their limits. However, few people have the time to study this vast body of literature in any detail, instead most numerical techniques are taken from pre-written software packages. The work in this thesis has been no exception, however it is worth reviewing the basis of these techniques. Such a review helps to clarify what methods have been used and more importantly, what their limitations are.

First, the problem of finding the roots of an equation will be discussed. This is a fundamental technique in the numerical solution of equations. The solution of the semiconductor equations has provided the basis for Chapter 3, so here the Poisson equation is discussed, and a simple method for its solution is described. This is then followed by the example of solving the 1-D semiconductor equations numerically. The generalisation to 2-D is far from trivial, but the 1-D case illustrated the techniques used.

Finding the Roots of an Equation (The Newton-Raphson Method)

To understand how we might solve an equation of the form \( f(x) = 0 \) first consider the example function shown in Figure A-1. There are two values of \( x \) at which \( f(x) \) equals zero and these are our roots. Now let us make an initial guess at a value for one root, if we guess well, then the process will take less time. We shall denote the initial guess as \( x_0 \). Now we wish to improve on this guess. To do this note that the
tangent at $x_0$ is given by equation (A-1) below, where $x_1$ is the intercept of the tangent on the x axis.

$$f'(x_0) = \left( \frac{df(x)}{dx} \right)_{x_0} = \frac{f(x_0)}{x_0 - x_1}$$

(A-1)

This can be solved (see equation A-2) to give a value for $x_1$ which is a better estimate for the value of $x$ at which $f(x)=0$ (this assumes that the function is continuous and well behaved). Then the value $x_1$ can be put back into equation (A-2) to give a better fit. Iterating this procedure will give progressively better fits, until successive values are identical, within the desired margin of accuracy.

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

(A-2)

Clearly, this depends on the above equation converging to a value. Care must be taken to only carry out the iteration a fixed number of times to ensure that the computer does not get caught in a loop. In addition, for the case of multiple roots, an initial poor guess may result in a root that is not wanted or missing a root that is wanted. Another difficulty is that, should the gradient of the function be close to zero, then the iteration could shoot off to infinity never to return. However, taking note of these difficulties, the Newton-Raphson methods is a frequently used technique in the numerical solution of algebraic equations.
Figure A-1: An Illustration of the Newton Raphson Method.

There are other techniques for discovering the roots of equations. These can be vast improvements on the Newton-Raphson method in 1-D but for multi-dimensional problems the Newton-Raphson method remains one of the few usable techniques.

Consider the situation where there are several functions, which we shall denote by $F_i$, where $i$ is an index identifying each such function. We also identify the variables of the function by $x_j$ where $j$ is again an index identifying the variables. Assuming that there are $N$ such variables and equations, then we wish to find the roots, i.e. find those values of $x_j$ which will satisfy the condition given by equation (A-3). Expanding this equation using Taylor series leads us to equation (A-4), where the usual notation of bold text is used to indicate vectors of arbitrary dimension.

\[
F_i(x_1, x_2, \ldots, x_N) = 0 \quad \text{(A-3)}
\]

\[
F_i(x + \delta x) = F_i(x) + \sum_{j=1}^{N} \frac{\partial F_i}{\partial x_j} \delta x_j + O(\delta x^2) \quad \text{(A-4)}
\]
By neglecting quadratic and higher terms and recognising the summation as the Jacobian, J, a set of linear equations can be formed. These can be solved using matrix techniques to give a value for δx.

**Solving Poisson's equation**

Laplace’s equation is a partial differential equation and many of the techniques applied to its solution can be used with other partial differential equations. As an example of the solution of Laplace’s equation, we shall use the two dimensional form given in equation (A-5) below. We shall then derive what is known as the finite difference form of this equation, this can be easily generalised to 3 or more dimensions at the cost of computing power.

\[
\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0 \tag{A-5}
\]

To derive the finite difference formula, it is necessary to divide the space across which the solution is required into some sort of mesh. For simplicity, we shall use a square mesh (a grid) with a spacing of h between points, as shown in Figure A-2. The following arguments can however be generalised to different shapes of mesh and different spacings. In fact a triangular mesh is usually preferred because it can cover unusually shaped objects and is more flexible than a square mesh.
Figure A-2: Example mesh for the solution of equation (A-5).

Using the Taylor’s theorem we can expand the potential about the point \((i,j)\) to give the potential at say \((i+1,j)\). Noting that \(\delta x\) equals \(h\) if \(h\) is small, then we can rearrange equation (A-5) to give equations (A-6) and (A-7).

\[
V_{i+1,j} = V_{i,j} + \left(\frac{\partial V}{\partial x}\right)_{i,j} \delta x
\]

(A-6)

\[
\left(\frac{\partial V}{\partial x}\right)_{i,j} = \frac{V_{i+1,j} - V_{i,j}}{h}
\]

(A-7)

Expanding the first derivative about the point \((i,j)\) using similar arguments, gives the second derivative as shown in equation (A-8).

\[
\left(\frac{\partial^2 V}{\partial x^2}\right)_{i,j} = \frac{1}{2h^2}(V_{i+1,j} + V_{i-1,j} - 2V_{i,j})
\]

(A-8)

Again using the same methods, i.e. expanding using a Taylor series and rearranging, gives equation (A-9) as an equivalent form of the Laplace equation.
Hence, the potential at a point in the grid is expressed in terms of the potential of its neighbours. Clearly the points that lie on the edge of the grid must have fixed values, these are the boundary conditions. When solving the potential in a semiconductor device this condition could be the bias at the contacts or zero at an insulating interface. Possibly, the device may have periodic boundary conditions. That is, it is assumed that another identical device is in contact with the simulated device and the boundary values are calculated by ‘wrapping the grid around’. So values on the right of the grid are calculated as if the values on the left of the grid are adjacent to them and vice versa.

By calculating the values at each point in the grid using equation (A-9) and repeating the process until the values no longer change, a solution can be obtained. It follows from the above, that the accuracy of the calculation depends upon the mesh, which also determines the resolution of the result.

**Solving the Semiconductor Current Relations**

To develop a numerically stable technique for solving the semiconductor equations is far from a trivial problem and some hint at the bewildering array of techniques can be found in the books by Carey et al. and Selberherr. Here we outline the Scharfetter and Gummel scheme, the first numerical method applied to semiconductor devices.
This scheme begins by scaling the semiconductor equations to give the equations below.

\[ \lambda^2 \left( \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right) - (n - p - C) = 0 \quad (A-10) \]

\[ \frac{\partial}{\partial y} \left( D_n \frac{\partial n}{\partial x} - \mu_n n \frac{\partial \psi}{\partial x} \right) + \frac{\partial}{\partial y} \left( D_n \frac{\partial n}{\partial y} - \mu_n n \frac{\partial \psi}{\partial y} \right) - R(n, p, \psi) = 0 \quad (A-11) \]

\[ \frac{\partial}{\partial y} \left( D_p \frac{\partial p}{\partial x} - \mu_p p \frac{\partial \psi}{\partial x} \right) + \frac{\partial}{\partial y} \left( D_p \frac{\partial p}{\partial y} - \mu_p p \frac{\partial \psi}{\partial y} \right) - R(n, p, \psi) = 0 \quad (A-12) \]

Here \( \lambda^2 = \psi_0 \xi/x_0^2 qC_0 \) is a scaling factor. The net doping concentration, \( C \), and the carrier concentrations \( n \) and \( p \) are scaled by \( C_0 \), the maximum doping. The potential \( \psi \), is scaled by the thermal voltage (\( k_B T/q \)). The diffusion coefficients \( (D_n \text{ and } D_p) \) are scaled by the maximum diffusion coefficient and \( x_0 \) is the maximum device dimension, which scales all spatial dimensions. The Recombination/Generation function, \( R \), depends upon the scaled variables and hence it is automatically scaled.

Following on from the solution of the Poisson equation, it is logical to use the same finite difference formulae here. To keep the calculations short and simple, only the electron equations will be considered. The hole equations can be derived in an identical manner; similarly, we shall neglect the \( y \) dimension. Thus, the finite difference form for equation (A-11) is given below in equation (A-13). Here the notation \( O(\delta) \) indicates terms containing the variable \( x \) which are neglected.

\[ \frac{J_{n|i+1/2} - J_{n|i-1/2}}{(h_x + h_{i-1})/2} + O(h) \frac{\partial^2}{\partial x^2} J_{n|i} - R(\psi, n, p) = 0 \quad (A-13) \]

\[ J_n = \mu_n n \frac{\partial \psi}{\partial x} - D_n \frac{\partial n}{\partial x} \quad (A-14) \]
An Introduction to Numerical Methods

We note that by carrying out a Taylor series expansion of the electron current across such a mesh, the finite difference form of the current is given by equation (A-15). Note that this equation gives the current density at any point \( x \) that lies between the mesh point \( i \) and \( i+1 \), and is not restricted to the mesh points.

\[
J_n(x) = J_n|_{i+1/2} + \left( x - x_i - \frac{h_i}{2} \right) \frac{\partial}{\partial x} J_n|_{i+1/2} + O(h^2) \frac{\partial^2}{\partial x^2} J_n|_{i+1/2}
\]  

(A-15)

Substituting equation (A-15) into equation (A-14) gives partial differential equation (A-16) shown below. The boundary conditions for this equation are set by the mesh points such that \( n(x_i) = n_i \) and \( n(x_{i+1}) = n_{i+1} \), i.e. the discrete solution and the continuous solution must match at the nodes of the mesh.

\[
J_n|_{i+1/2} + \left( x - x_i - \frac{h_i}{2} \right) \frac{\partial}{\partial x} J_n|_{i+1/2} = \mu_n \frac{\partial n}{\partial x} - D_n \frac{\partial n}{\partial x}
\]  

(A-16)

A solution to this equation can be found by if it is assumed that there is a constant electric field between any two neighbouring mesh points. Thus, the carrier concentration between nodes of the mesh is given by equation (A-17).

\[
n(x) = (1 - g_i(x, \psi))n_i + g_i(x, \psi)n_{i+1}
\]  

(A-17)

\[
g_i(x, \psi) = \frac{1 - \exp \left( \frac{\psi_{i+1} - \psi_i}{h_i} \right) \frac{x - x_i}{h_i}}{1 - \exp \left( (\psi_{i+1} - \psi_i) \right)}
\]  

(A-18)

Using (A-17) with (A-16) a solution for the current at a given node can be found. The results, when placed into (A-13), give the discretised continuity equation for electrons, (A-19). The form of this equation is intuitive, simply being a statement of the fact that the total current entering a mesh node must equal the current leaving that node. It is complicated by the exponential form that the carrier concentration takes.
\[
D_n|_{i+1/2} \frac{B(\psi_{i+1} - \psi_i)n_{i+1} - B(\psi_i - \psi_{i+2})n_i}{h_i \frac{h_i + h_{i+1}}{2}} - D_n|_{i-1/2} \frac{B(\psi_i - \psi_{i-1})n_i - B(\psi_{i-1} - \psi_i)n_{i-1}}{h_{i-1} \frac{h_i + h_{i-1}}{2}} - R(\psi, n, p) = 0
\]  
(A-19)

Where the Bernoulli function has been introduced and is defined by equation (A-20).

\[
B(x) = \frac{x}{\exp(x) - 1}
\]  
(A-20)

This can be solved in conjunction with the Poisson equation and the Hole continuity by using a multi-dimensional form of the Newton-Raphson method.

**Evaluating Integrals**

To evaluate an integral is a simple process as is illustrated by Figure A-3. The integral of a function is simply the area under the function within the limits of the integral. To evaluate this area one simply takes two points on a curve and draws a straight line between them to generate a trapezoid. In the limit that the distance between the two points goes to a small value, then the area of the trapezoid approaches that under the curve.
Figure A-3: Illustration of a technique for numerical integration.

References


Appendix B Derivation of the Defect Pool Model and the Occupation Statistics

Occupation Statistics

To begin this derivation we refer to the work of Shockley and Last\(^1\) who derive occupancy statistics for defects of arbitrary charged states. Consider a defect, which has \(s\) electronic charges and a degeneracy described by \(i\). \(N_{si}\) is defined as the number of defects in excited state \(i\) with charge condition \(s\). So the total number of defects is the sum of all of these states and the number of electrons in the flaws are given by equation (B-1).

\[
E_{DB} = \sum_{s,i} (s-1)N_{si}
\]  
(B-1)

For this derivation, the positive charge state is assumed to have zero electrons to ensure the validity of the above equation. However the equation is quite general and can be extended to arbitrary systems with suitable modifications. It now follows that the energy of a given distribution of dangling bonds is given by equation (B-2).

\[
E_{DB} = \sum_{s,i} E_{si}N_{si}
\]  
(B-2)

To arrive at the occupancy statistics we must find the most probable way of forming the distribution \(N_{si}\). The number of ways of selecting distributions of \(N_{si}\) from \(M\) defects is given by equation (B-3).

\[
W_{DB} = \frac{M!}{\prod_{s,i} N_{si}!} = \left( \frac{\sum N_{si}}{\prod N_{si}} \right)! \]  
(B-3)
Derivation of the Defect Pool Model and the Occupation Statistics

To account for the possibility of electrons going into the conduction band, we can also find the most probable way of arranging electrons into a density of states $g_i$ at and energy $E_j$ with each group of states containing $n_i$ electrons. This gives equations (B-4) to (B-6) below describing the conduction band distributions, their total number and energies respectively.

$$ W_{cb} = \prod_j \frac{g_j}{n_j!(g_j - n_j)!} \quad (B-4) $$

$$ n_{cb} = \sum_j n_j \quad (B-5) $$

$$ E_{cb} = \sum_j E_j n_j \quad (B-6) $$

From the basic postulates of statistical mechanics the most probable distribution should be given by maximising equations (B-3) and (B-4) subject to the constraints that the total number of electrons and defects remains constant, also that the energy is conserved. This can be solved using the method of Lagrangian multipliers to give equation (B-7) below.

$$ \ln W_{DB} + \ln W_{cb} - \beta(E_{DB} + E_n) + \alpha(n_{DB} + n_{cb}) + \gamma M $$

$$ = \ln \left( \frac{M!}{\prod_s N_{s!}} \right) + \ln \left( \prod_j \frac{g_j!}{n_j!(g_j - n_j)!} \right) - \beta(E_{DB} + E_n) + \alpha(n_{DB} + n_{cb}) + \gamma M \quad (B-7) $$

Using the properties of logarithms, i.e. that $\ln (A/B) = \ln (A-B)$, we can simplify the right hand side of the above equation to equation (B-8) below.

$$ \ln M! - \sum_{s,l} \left( \ln N_{s!} + \sum_j \left( \ln (g_j!) - \ln (n_j!) - (g_j - n_j)! \ln ((g_j - n_j)!) \right) \right) $$

$$ - \beta(E_{DB} + E_n) + \alpha(n_{DB} + n_{cb}) + \gamma M \quad (B-8) $$
Differentiating with respect to $N_{si}$ and $n_j$ leads to the following equation for the most probable distribution.

$$\frac{\partial}{\partial N_{si}} \left( \ln M! - \sum_{s,i} (\ln N_{si}) - \beta E_{DB} + \alpha n_{DB} + \gamma M \right) dN_{si} +$$

$$\frac{\partial}{\partial n_j} \left( \sum_{j} \left( \ln(g_j!) - \ln(n_j!) - (g_j - n_j)! \ln((g_j - n_j)!) \right) - \beta E_{n} + \alpha n_{c} \right) dn_j = 0 \tag{B-9}$$

Choosing the case where the coefficient of $dn_j = 0$, gives the Fermi-Dirac case for the conduction band electrons. Selecting the coefficient of $dN_{si} = 0$ leads to the equation (B-10) below. For equation (B-10) Stirling’s approximation, $\ln A! \approx A \ln A - A$, has been used to simplify the expression and we note that $M$ is defined as being the sum across all states of $N_{si}$.

$$\frac{\partial}{\partial N_{si}} \left( \sum_{i,s} N_{si} \ln \left( \sum_{i,s} N_{si} \right) - \sum_{i,s} N_{si} \right) - \sum_{i,j} \left[ N_{si} \ln N_{si} - N_{si} + \beta N_{si} E_{si} - \alpha (s - 1) N_{si} - \gamma N_{si} \right] = 0 \tag{B-10}$$

If we note that by a suitable choice of the multipliers, this equation can be made true for all possible states, then we can separate all of the summations into separate equations and consider only a single general case which reduces to:-

$$1 - \ln M - 1 - (1 - \ln N_{si} - 1 + \beta E_{se} - \alpha (s - 1) - \gamma) = 0$$

$$\ln \left( \frac{M}{N_{si}} \right) - \beta E_{se} + \alpha (s - 1) + \gamma = 0 \tag{B-11}$$

And this leads directly to the equation giving the number of centres in charge state $s$ as being:

$$N_s = \sum_i N_{si} = M \exp(\gamma + \alpha (s - 1)) \sum_i \exp(-\beta E_{si}) \tag{B-12}$$
Before proceeding further we note that for our case we wish to consider the case of a
dangling bond which has a charge occupancy of +e, 0 or −e. This means that \( s = -1, 0, 
1 \) and that we can define the energies \( E_{si} \) as shown in equations (B-13) to (B-15).
Here \( E \) is the energy of one electron in a defect and \( U \) is the correlation energy
introduced by bringing two electrons together in the same state.

\[
E_{-1} = 2E + U \quad \text{being the energy of the } D^- \text{ state} \quad \text{ (B-13)}
\]
\[
E_0 = E \quad \text{being the energy of the } D^0 \text{ state} \quad \text{ (B-14)}
\]
\[
E_{+1} = 0 \quad \text{being the energy of the } D^+ \text{ state} \quad \text{ (B-15)}
\]

The undetermined multipliers can now be evaluated. We begin by noting that \( \beta \) is in
fact the statistical temperature\(^2\) and is the inverse of \( k_bT \), where \( T \) is temperature in
Kelvin and \( k_b \) is Boltzmann’s constant. In a similar manner, we can associate \( \alpha \) with
the chemical potential or Fermi level. This is a definition but has a sound basis in the
transition from statistical mechanics to thermodynamics, where \( \alpha \) is found to relate to
the chemical potential through the relationship \( \alpha = \beta E_f \) where \( E_f \) is the Fermi energy.

As \( \gamma \) was introduced to maintain the condition of constant defect density we can
normalise the equation to ensure that this condition is fulfilled such that:

\[
\sum_u (N_u e^\gamma)^{\frac{1}{e^\gamma} - 1} = M e^\gamma \
\]

\[
\sum_u \left[ \exp\left( \frac{E_f}{k_bT} (1-1) \right) + 2 \exp\left( \frac{[0-1](E_f - E)}{k_bT} \right) \right] = M \quad \text{(B-16)}
\]

This leads us directly to the partition function for the system such that:-

\[
\exp(\gamma) = \frac{1}{1 + 2 \exp\left( \frac{E_f - E}{k_bT} \right) + \exp\left( \frac{2E_f - 2E - U}{k_bT} \right)} \quad \text{(B-17)}
\]
We can now write out the occupancy functions for each of the charge states:

\[ f^+(E) = \frac{1}{1 + 2 \exp\left(\frac{E_F - E}{k_B T}\right) + \exp\left(\frac{2E_F - 2E - U}{k_B T}\right)} \]  
\[ f^0(E) = \frac{2 \exp\left(\frac{E_F - E}{k_B T}\right)}{1 + 2 \exp\left(\frac{E_F - E}{k_B T}\right) + \exp\left(\frac{2E_F - 2E - U}{k_B T}\right)} \]  
\[ f^-(E) = \frac{\exp\left(\frac{-2E_F - 2E - U}{k_B T}\right)}{1 + 2 \exp\left(\frac{E_F - E}{k_B T}\right) + \exp\left(\frac{2E_F - 2E - U}{k_B T}\right)} \]  

**The Improved Defect Pool Model**

The defect pool model begins by using the above equations to derive the following expressions for the average energy of any electrons in these defect states, \(\langle e \rangle\), and the entropy, \(s_e\).

\[ \langle e \rangle = E_F f^+(E) + Ef^0(E) + [2E - E_F + U]f^-(E) \]  
\[ s_e = -\sum_{\text{all states}} p_i \ln p_i = \left\{ f^+(E) \ln(f^+(E)) + 2\left[ \frac{f^0(E)}{2} \right] \ln\left( \frac{f^0(E)}{2} \right) + f^-(E) \ln(f^-(E)) \right\} \]
Derivation of the Defect Pool Model and the Occupation Statistics

where $p_i$ is the probability of the system being in a given state, $i$.

From the definition of chemical potential, $\mu_d$, and using the occupancy functions, the above equations can be used to arrive at the expression for chemical potential given in equation (B-23).

$$\mu_d = \langle e \rangle - k_b T s$$

$$= 2E - E_p + U + k_b T \ln f^-(E) \quad \text{(B-23)}$$

In the above calculation, the assumption is made that the defect density is constant. However, the defect pool model adds the assumption that the action of hydrogen in the material can create and destroy defects by breaking weak Si-Si bonds. The improved nature of Powell and Deane's model is the use of three equations to describe the possible defect generating reactions instead of the usual two. The reactions are given below.

$$\text{WB} \leftrightarrow (2\text{D}) \quad \text{(B-24)}$$

$$\text{Si-H} + \text{WB} \leftrightarrow (\text{D} + \text{Si-H}) + \text{D} \quad \text{(B-25)}$$

$$2\text{Si-H} + \text{WB} \leftrightarrow (\text{Si-H} - \text{H} - \text{Si}) + 2\text{D} \quad \text{(B-26)}$$

where WB indicates a weak bond.

If we choose $i$ to be the number of Si-H bonds involved in any chemical interaction, then we can arrive at the following expression for the probability that a defect exists at a given hydrogen site, $p_d$.

$$p_d = \frac{iD(E)}{2HP(E)} \quad \text{(B-27)}$$
Derivation of the Defect Pool Model and the Occupation Statistics

where \( D(E) \) is the density of defects at a given energy, \( E \). \( H \) is the density of hydrogen and \( P(E) \) is the distribution of sites that may produce defects (i.e. the Si-Si weak bond density).

The origin of this expression can be seen by considering that there are \( D(E) \) defects at any given time from a possible selection of \( H \times P(E) \) defect sites. However, the \( i/2 \) term is required as for every two defects formed only \( i \) defects can swap hydrogen with a distant site.

Repeating the earlier process used to determine the chemical potential but including the additional term for entropy arising from the hydrogen interactions gives a defect chemical potential of:

\[
\mu_d(E) = E + k_b T \left[ \frac{f^0(E)}{2} + \frac{i k_b T}{2} \ln \left( \frac{iD(E)}{2HP(E)} \right) \right]
\]  
(B-28)

Now consider the transition from a weak bond with an energy \( E_t \) into a defect site with an energy of \( E \). If there are \( P(E)g_t(E_t) \) weak bonds that have not been converted to defects and \( D(E) \) is the density of defect states, then by treating these as a two level thermodynamic system we arrive at equation (B-29), below. The factor of 2 in the exponent arises because for each weak bond broken, two defects are created.

\[
D(E) = \left[ P(E)g_t(E_t) - D(E) \right] \exp \left[ \frac{-2(\mu_d(E) - E_t)}{k_b T} \right]
\]  
(B-29)

The weak bond states are identified with the valence band tail states, which are expected to have a density of states of the form given by equation (B-30).

\[
g_t(E_t) = N_{v0} \exp \left( \frac{(E_t - E_{v0})}{E_{v0}} \right)
\]  
(B-30)
Derivation of the Defect Pool Model and the Occupation Statistics

where \( N_{v0} \) is the density of tail states at the mobility edge, \( E_v \) is the valance band energy and \( E_{v0} \) is the characteristic energy of the distribution. By assuming that those states below the chemical potential convert, whilst those above it convert with a probability given by the Boltzmann distribution, the above equation can be integrated simply. If the defect pool of weak bond states is assumed to be given by equation (B-31) below, we can then carry out this integration to arrive at equation (B-32).

\[
P(E) = \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{E-E_p}{2\sigma^2} \right]
\]

where \( \sigma \) is the width of the defect pool density and \( E_p \) is the most probable energy for a defect.

\[
D(E) = \gamma \frac{2}{f^0(E)} \frac{\sigma^2}{k_B T} \exp \left[ -\frac{E_p}{E_{v0}} \right] \left[ E + \frac{\rho \sigma^2}{2E_{v0}} \right]
\]

and the following abbreviations have been introduced:

\[
\gamma = \left( \frac{N_{v0} 2E_{v0}^2}{(2E_{v0} - k_B T)} \right) \frac{1}{2H} \frac{1}{\rho} \exp \left[ -\frac{E_p}{E_{v0}} \left( E_v - E_p - \frac{\rho \sigma^2}{2E_{v0}} \right) \right]
\]

\[
\rho = \frac{2E_{v0}}{(2E_{v0} + k_B T)}
\]

This is the principle equation of the defect pool model as discussed in chapter 4. From this equation, we can calculate an appropriate density of states, which can be broken up into one electron density of states using the following relations.

\[
D^+ (E,T) = f^+ (E,T) D(E,T_e)
\]
\[
D^0 (E,T) = f^0 (E,T) D(E,T_e)
\]
\[
D^- (E,T) = f^- (E,T) D(E,T_e)
\]

(B-35)
Here we have used $T$ to indicate the temperature dependence and $T_e$ to mark an equilibration temperature. What this means is that below a certain temperature the density of states is assumed to be 'frozen' in so that the defect density no longer changes. Consequentially, when evaluating $D(E)$, if the temperature is below the equilibration temperature, $T_e$, then the equilibration temperature should be used otherwise the actual temperature is used. Note that this does not affect the occupancy functions, which are always expected to use the actual temperature.

References


Appendix C Capture and Emission coefficients in Dangling Bond recombination

In chapter 4, the dangling bond equations were expressed in terms of capture and emission coefficients. Equations for these coefficients are given in Table C-1 below. These equations come from the assumption that the capture process can be described by a simple cross section. Hence, from the definition of cross section, the probability of a carrier being captured is the product of the thermal velocity of that carrier and the cross section for capture. To assess the rate of capture this must be multiplied by the number of free carriers and the number of trapping states.

To arrive at the emission coefficients it is assumed that the carriers are hopping from defect states to band states using phonons to gain sufficient energy for the leap. Such a process is detailed by Mott and Davis\(^1\) and results in the relationship:

\[
e = \omega \exp\left(\frac{-E_a}{k_b T}\right) \quad (C-1)
\]

Where \(E_a\) is the activation energy for the process and \(\omega\) is the maximum phonon frequency.

The activation energy is the difference between the energy of the defect state and the band tail for one carrier interactions. However, for \(D\) states the initial energy is \(2E_d-U\) and the final energy is \(E_c + E_d\), leading to an activation energy of \(E_c - E_d - U\). These considerations lead us to the values given in Table C-1.
Capture and Emission coefficients in Dangling Bond recombination

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Neutral States</th>
<th>Negative State</th>
<th>Positive States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electron Emission, $e_n$</td>
<td>$\omega \exp \left( -\frac{E_c - E_d}{k_b T} \right)$</td>
<td>$\omega \exp \left( -\frac{E_c - E_d - U}{k_b T} \right)$</td>
<td>0</td>
</tr>
<tr>
<td>Electron Capture, $e_n$</td>
<td>$V_n \sigma_n^0(E_d)$</td>
<td>$V_n \sigma_n^-(E_d)$</td>
<td>$V_n \sigma_n^+(E_d)$</td>
</tr>
<tr>
<td>Hole Emission, $e_p$</td>
<td>$\omega \exp \left( \frac{E_v - E_d}{k_b T} \right)$</td>
<td>0</td>
<td>$\omega \exp \left( \frac{E_v - E_d}{k_b T} \right)$</td>
</tr>
<tr>
<td>Hole Capture, $e_p$</td>
<td>$V_p \sigma_p^0(E_d)$</td>
<td>$V_p \sigma_p^-(E_d)$</td>
<td>$V_p \sigma_p^+(E_d)$</td>
</tr>
</tbody>
</table>

In the above the superscripts denote the charge state to which that parameter is linked; 0 indicates a neutral state, + is a positively charged state and - is a negative state. The subscripts p and n mark the carrier as either holes or electrons respectively. V is the thermal velocity of the marked carrier, $k_b$ is the Boltzmann’s constant (eV K⁻¹), T is the temperature (K), $\omega$ is the maximum phonon frequency (Hz), $E_d$ is the energy of the defect state, $E_c$ is the conduction band energy and $E_v$ is the valence band energy.

Table C-1: Equations for the various capture coefficients used in calculating the rate equations in chapter 4.

These coefficients provide the basis for the analysis of chapter 4; however it is worth mentioning the work of Kramer and van Berkel. They point out that in addition to the above terms there is the possibility of tunnelling through virtual states close to the conduction band edge which adds a further term to the emission coefficients. More interestingly, they propose that an additional term should be included in the emission coefficient term to account for the possibility of Poole-Frenkel tunnelling. This
Capture and Emission coefficients in Dangling Bond recombination means that the activation energies in Table C-1 should be lowered by $2(eF/4\pi\varepsilon_0\varepsilon_r)^{1/2}$ when the emission occurs from neutral traps, where $e$ is the charge on an electron, $F$ is the electric field, $\varepsilon_0$ is the permittivity of free space and $\varepsilon_r$ is the permittivity of the material.

Such a term may account for the increase in recombination with increasing reverse bias voltage. From the literature, typical values for the capture coefficients are estimated at around $1 \times 10^{-15}$ cm$^{-2}$ for charged states and at about $1 \times 10^{-17}$ cm$^{-2}$ for neutral states. The maximum phonon frequency is expected to have a value of about $10^{12}$ Hz.

Table C-2 shows the values that might be expected for these coefficients assuming a band gap 1.9 eV and a temperature of 300 K. From these values, it can be seen that scaling the coefficients is desirable. A factor of $1 \times 10^{-10}$ gives an effective range of 0.01 to 100 with values much lower than that being reduced to insignificance, a useful range for numerical calculations.
<table>
<thead>
<tr>
<th>Energy</th>
<th>$e_n^0$</th>
<th>$e_n^+$</th>
<th>$e_p^0$</th>
<th>$e_p^-$</th>
<th>$e_n^0$</th>
<th>$e_n^-$</th>
<th>$e_p^0$</th>
<th>$e_p^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.4E-20</td>
<td>3.1E-17</td>
<td>3.1E+12</td>
<td>1.0E+12</td>
</tr>
<tr>
<td>0.1</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>6.6E-19</td>
<td>1.5E-15</td>
<td>2.1E+10</td>
<td>2.1E+10</td>
</tr>
<tr>
<td>0.2</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>3.1E-17</td>
<td>7.0E-14</td>
<td>4.4E+08</td>
<td>4.4E+08</td>
</tr>
<tr>
<td>0.3</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.5E-15</td>
<td>3.3E-12</td>
<td>9.3E+06</td>
<td>9.3E+06</td>
</tr>
<tr>
<td>0.4</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>7.0E-14</td>
<td>1.6E-10</td>
<td>2.0E+05</td>
<td>2.0E+05</td>
</tr>
<tr>
<td>0.5</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>3.3E-12</td>
<td>7.6E-09</td>
<td>4.1E+03</td>
<td>4.1E+03</td>
</tr>
<tr>
<td>0.6</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.6E-10</td>
<td>3.6E-07</td>
<td>8.7E+01</td>
<td>8.7E+01</td>
</tr>
<tr>
<td>0.7</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>7.6E-09</td>
<td>1.7E-05</td>
<td>1.8E+00</td>
<td>1.8E+00</td>
</tr>
<tr>
<td>0.8</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>3.6E-07</td>
<td>8.1E-04</td>
<td>3.9E-02</td>
<td>3.9E-02</td>
</tr>
<tr>
<td>0.9</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.7E-05</td>
<td>3.9E-02</td>
<td>8.1E-04</td>
<td>8.1E-04</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>8.1E-04</td>
<td>1.8E+00</td>
<td>1.7E-05</td>
<td>1.7E-05</td>
</tr>
<tr>
<td>1.1</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>3.9E-02</td>
<td>8.7E+01</td>
<td>3.6E-07</td>
<td>3.6E-07</td>
</tr>
<tr>
<td>1.2</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.8E+00</td>
<td>4.1E+03</td>
<td>7.6E-09</td>
<td>7.6E-09</td>
</tr>
<tr>
<td>1.3</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>8.7E+01</td>
<td>2.0E+05</td>
<td>1.6E-10</td>
<td>1.6E-10</td>
</tr>
<tr>
<td>1.4</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>4.1E+03</td>
<td>9.3E+06</td>
<td>3.3E-12</td>
<td>3.3E-12</td>
</tr>
<tr>
<td>1.5</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>2.0E+05</td>
<td>4.4E+08</td>
<td>7.0E-14</td>
<td>7.0E-14</td>
</tr>
<tr>
<td>1.6</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>9.3E-06</td>
<td>2.1E+10</td>
<td>1.5E-15</td>
<td>1.5E-15</td>
</tr>
<tr>
<td>1.7</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>4.4E-08</td>
<td>1.6E+12</td>
<td>3.1E-17</td>
<td>3.1E-17</td>
</tr>
<tr>
<td>1.8</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>2.1E+10</td>
<td>4.8E+13</td>
<td>6.6E-19</td>
<td>6.6E-19</td>
</tr>
<tr>
<td>1.9</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E-10</td>
<td>1.0E-08</td>
<td>1.0E+12</td>
<td>2.3E+15</td>
<td>1.4E-20</td>
<td>1.4E-20</td>
</tr>
</tbody>
</table>

Table C-2: Typical values for capture and emission coefficients.

References


Appendix D  Silvaco Software

D.1 A Background to the Silvaco Software

The semiconductor simulations in the thesis have been based on the semiconductor simulation software supplied by Silvaco Software. The basis for this software is the Stanford simulators Pisces (for device simulation) and SUPREM IV (for process simulation). Also used is the devedit mesh generation software developed by Integrated Systems Laboratory. This software has been developed and expanded by Silvaco to include, among other things, amorphous silicon recombination models.

The simulators are worked by writing a ‘deck’, which describes the conditions to be simulated. To serve as an example of the work that has been carried out in chapter 3, the next sections give annotated versions of decks used to solve two problems. The first in section D.2 is a devedit deck, which constructs a simple photodiode (the diode used in Chapter 3 for the simulation ID#2). Following on from this in section D.3 is a deck, which simulates the reverse and forward bias characteristics of a diode and then illuminates it under reverse bias conditions.

D.2 Sample Silvaco Deck for Devedit: Device/Mesh Generation

```
go devedit
#These set variables to define the device
#parameters
set BASE_LINE = 2.0
set SUBSTRATE_WIDTH = 100
set SUBSTRATE_DEPTH = 0.5
set ANODE_WIDTH = 60
set ANODE_THICKNESS = 0.1
set CATHODE_WIDTH = $SUBSTRATE_WIDTH
```
set CATHODE_THICKNESS = 0.1
set INTRINSIC_THICKNESS = 1.0
set ANODE_DOPING = 1e15
set ANODE_REF = 1e10
set ANODE_OVERLAP = 2.0
set ANODE_DEPTH = 0.1
set CATHODE_DOPING = 1e15
set CATHODE_REF = 1e10
set CATHODE_DEPTH = 0.1

# This defines a substrate region.
# (X1,Y1) and (X2,Y2) define the corners of the
# substrate.
# Substrate (Reg. 1)
set X1 = 0
set X2 = $SUBSTRATE_WIDTH / 2
set Y1 = $BASE_LINE
set Y2 = $Y1 - $SUBSTRATE_DEPTH
set POLY_DEFINE = "$X1 , $Y1 $X1 , $Y2 $X2 , $Y2 $X2 , $Y1 $X1 , $Y1"

# This defines the region number and material
region reg=1 mat="Silicon Oxide" color=0xff \ 
pattern=0x2 polygon=" $POLY_DEFINE "

# This defines the first electrode.
# Anode (Reg. 2; Elec. 1)
set X1 = 0
set X2 = $ANODE_WIDTH / 2
set Y1 = $BASE_LINE - $SUBSTRATE_DEPTH
set Y2 = $Y1 - $ANODE_THICKNESS
set POLY_DEFINE = "$X1 , $Y1 $X1 , $Y2 $X2 , $Y2 $X2 , $Y1 $X1 , $Y1"

# Unlike the substrate electrodes must
# have an id number and a name for
# identification in the device simulator
region reg=2 name=anode mat=Aluminum \ 
elec.id=1 color=0xffc8c8 pattern=0x7 \ 
polygon=" $POLY_DEFINE "
#Here the Silicon layer is defined.

**Intrinsic Layer (Reg. 3)**

set $X_1 = 0$

set $X_2 = \frac{ANODE\_WIDTH}{2}$

set $X_3 = \frac{SUBSTRATE\_WIDTH}{2}$

set $Y_1 = BASE\_LINE - SUBSTRATE\_DEPTH$

set $Y_2 = Y_1 - ANODE\_THICKNESS$

set $Y_3 = Y_2 - INTRINSIC\_THICKNESS$

set \( POLY\_DEFINE = \"X_1, Y_2, X_2, Y_2, X_2, Y_1, X_3, Y_1, X_3, Y_3, X_1, Y_3, X_1, Y_2 \\)$

region reg=3 mat=Silicon color=0xffcc00 pattern=0x4 \ polygon=" POLY\_DEFINE "

**Intrinsic Doping** - This defines a background doping of phosphorous throughout the Silicon region.

impurity id=1 region.id=3 imp=Phosphorus \ peak.value=10000000000 comb.func=Multiply

**Cathode** This defines the cathode

set $X_1 = 0$

set $X_2 = \frac{CATHODE\_WIDTH}{2}$

set $Y_1 = BASE\_LINE - SUBSTRATE\_DEPTH - ANODE\_THICKNESS - INTRINSIC\_THICKNESS$

set $Y_2 = Y_1 - ANODE\_THICKNESS$

set \( POLY\_DEFINE = \"X_1, Y_1, X_1, Y_2, X_2, Y_2, X_2, Y_1, X_1, Y_1 \\)$

region reg=4 name=cathode mat=Aluminum elec.id=2 \ color=0xffc8c8 pattern=0x7 polygon=" POLY\_DEFINE "

**Anode Doping** - this defines the Anode doping as Boron with a Gaussian Profile.

set $X_1 = 0$

set $X_2 = \frac{ANODE\_WIDTH}{2} + ANODE\_OVERLAP$

set $Y_1 = BASE\_LINE - SUBSTRATE\_DEPTH$

set $Y_2 = Y_1 - ANODE\_THICKNESS - ANODE\_DEPTH$

impurity id=1 imp=Boron color=0x8c5d00 \ peak.value= ANODE\_DOPING ref.value=ANODE\_REF comb.func=Interpolate \ y1=$Y_2$ y2=$Y_1$ rolloff.y=low conc.func.y="Gaussian (Dist)" conc.param.y=0.1 \ x1=$X_1$ x2=$X_2$ rolloff.x=both conc.func.x="Gaussian (Dist)" conc.param.x=0.1
# This defines the doping around the cathode as phosphorous with a gaussian profile.
set X1 = 0
set X2 = $CATHODE_WIDTH
set Y1 = $BASE_LINE - $SUBSTRATE_DEPTH - $ANODE_THICKNESS - $INTRINSIC_THICKNESS
set Y2 = $Y1 - $CATHODE_THICKNESS
set Y1 = $Y1 + $CATHODE_DEPTH
impurity id=2 imp=Phosphorus color=0x8c5d00 \ peak.value= $CATHODE_DOPING ref.value= $CATHODE_REF comb.func=Interpolate \ yl= $Y2 y2= $Y1 rolloff.y=high conc.func.y="Gaussian (Dist)" conc.param.y=0.1 \ xl= $X1 x2= $X2 rolloff.x=both conc.func.x="Gaussian (Dist)" conc.param.x=0.1

# Set Meshing Parameters
# Defines a square mesh with y spacing of 0.25 microns and x spacing of 10 microns.
base.mesh height=0.25 width=10
# Conditions the boundaries between materials
# by preventing slopes greater than max.slope,
# joining lines that have < 1 degree difference
# in the slope and rounding boundary points to
# the nearest 1 nm.
bound.cond !apply max.slope=30 max.ratio=100 rnd.unit=0.001 line.straightening=1 align.points when=automatic
# Ensure that for every half decade change of the net doping there is at least 1 point.
imp.refine imp="Net Doping" scale=log sensitivity=0.5 transition=1e+12 imp.refine min.spacing=0.02
# Prevent the creation of mesh triangles which
# have angles > 90 degrees and height to width
# ratios greater than 100
constr.mesh max.angle=90 max.ratio=100 max.height=10000 \ max.width=10000 min.height=0.0001 min.width=0.0001

#Build the mesh
Mesh Mode=MeshBuild
D.3 Sample Silvaco deck for Atlas: Device Simulation

```
go atlas
# Load in the mesh from the devedit deck
mesh inf="Simple2.str"

# Define the models to use:-
# Solve the drift-diffusion equations at 300 K
# using Fermi-Dirac Statistics
models print fermi temperature=300 bbt.std drift.diff
# Modify the parameters to those more appropriate
# for c-Si. The f.index file reads in a c program
# to specify the optical constants.
material Material=Silicon eg300=1.72 \nc300=1e19 nv300=1e19 vthn=1e7 vthp=1e7 mun=10 \mup=1 f.index=rindex.lib
# Define an a-Si defect density. The f.tftdon
# and f.tftacc commands read in a c program that
# defines the defect density after Hack and Shur.
defects f.tftdon=hackl.lib f.tftacc=hackl.lib \NTA=0 NTD=0 WTA=1.0 WTD=1.0 NGA=0 NGD=0 \EGA=0.6 EGD=0.6 WGA=1 WGD=1 sigtah=1.e-16 \sigtae=1.e-14 sigtde=1.e-14 sigtdh=1.e-16 \siggae=1e-14 siggah=1e-16 siggde=1.e-14 \siggdh=1.e-16 \dfile=donors afile=acceptors

#Contact Definition - all contacts are ohmic.
contact name=anode neutral
contact name=cathode neutral

# Define a beam for an optical simulation
# The beam goes from 0 to 10 microns and has
# a wavelength of 550 nm
beam angle=90 min.window= 0 max.window= 10 \number=1 quantum.eff=1.0 reflects=0 thickest=1.0 \wavelength=0.550 x.origin= 5 y.origin=-1000
```
# Numerical Models Specification
# Use a Gummel method, switching to Newton-Raphson
# if convergence is not achieved. Setting climit
to 1e-4 means that errors in the carrier
# concentration must be less than this fraction of
# the current carrier concentration.
method newton autonr gummel carriers=2 climit=1e-4
trap atrap=0.25

# Solution Specification
# Find an initial solution to the poisson equation
solve init
# Solve the full set of drift-diffusion equations
# using the above solution as an initial guess.
solve prev
# Bias the anode contact to -10 V, stopping if
# the current exceeds 1 nA. The IV characteristics
# are logged in sim2rev.log
solve vanode=0.0
solve vanode=-0.1
log outf=sim2rev.log
solve name=anode vanode=-0.1 vstep=-0.1 vfinal=-10.0
cname=anode compliance=-1e-9 project

# Solve to find the optical response by applying
# a beam of increasing power. The results are
# logged to beaml.log
log outf=beaml.log
solve bl=1e-9
solve bl=1e-6
solve bl=1e-3
solve bl=1
solve bl=1e3
solve bl=1e6