ABSTRACT

PETERS, JEREMY. Intragrain Defect Characterization Of Solar Grade Silicon Using Near-Field Scanning Optical Microscopy. (Under the direction of Dr. George Rozgonyi.)

Multicrystalline silicon (mc-Si) is a material used in the photovoltaic (PV) industry because of its lower production cost in comparison to its single crystal or thin film silicon counterparts. Multicrystalline silicon grown by the block casting technique, in which molten silicon is cooled in a growth crucible, generates thermally induced stress, creating intragrain dislocation clusters. These dislocation clusters act as minority carrier recombination centers, reducing the overall cell efficiency. To gain understanding of the recombination behavior of these defects a characterization tool with submicron resolution is needed.

Materials scientists have a number of microscopy options at their disposal to characterize the structural, chemical, and electrical properties of semiconductors. Optical microscopy, e.g., differential interference contrast (DIC) techniques such as Nomarski microscopy, is used to observe surface defects delineated by chemical etching. However, far-field optical spatial resolution is limited by the Abbe limit, which states that the minimum resolvable distance between two objects is limited by the wavelength of the incident radiation. Electron microscopy, including scanning electron microscopy (SEM) and transmission electron microscopy (TEM), is employed to provide electrical and structural information about bulk defect interactions. Although capable of sub-angstrom resolution, electron microscopy requires sample preparation that destroys the sample surface. Advances in scanning probe microscopy (SPM) have allowed scientists to break the far-field limit to produce images with nanometer and subnanometer resolution.
Near-field scanning optical microscopy (NSOM) is a form of scanning probe microscopy. NSOM has the ability to image both surface and bulk properties of a material in a non-evasive manner with greater spatial resolution than far-field optical microscopy and electron microscopy. I intend to demonstrate NSOM as a characterization tool in photovoltaic (PV) silicon wafers, using carrier lifetime and photoinduced current variation as contrast mechanisms.

The motivation for and development of NSOM as an characterization tool to map defect recombination behavior is first described. Then, an account of the carrier dynamics associated with the NSOM contrast modes is given. Next, the design challenges associated with the construction of the NSOM system are explained. An analysis of the intragrain defect lifetime and recombination behavior follows, using results from both existing characterization techniques and NSOM imaging. Finally, a summary of findings and description of areas for future study is given.
Intragrain Defect Characterization Of Solar Grade Silicon

Using Near-Field Scanning Optical Microscopy

JEREMY KELVIN PETERS

A thesis submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
Requirements for the Degree of
Master of Science

MATERIALS SCIENCE AND ENGINEERING
RALEIGH, NORTH CAROLINA

2005

APPROVED BY:

Dr. George A. Rozgonyi
Material Science & Engineering
Co-Chair of Advisory Committee

Dr. Hans Hallen
Department of Physics
Co-Chair of Advisory Committee

Dr. Gerd Duscher
Materials Science & Engineering
Member of Advisory Committee
Dedicated to my parents, Melvin and Joy, and my Aunt Rose, who believed in me even in those times when I did not.
BIOGRAPHY

Jeremy Peters was born on April 8, 1976 in Worcester, Massachusetts. He graduated from Riverside High School in Durham, NC in 1995. He received his Bachelor’s Degree in Materials Science from Brown University in 2001. In February 2002, he worked as an undergraduate assistant in Dr. George Rozgonyi’s Microelectronic Materials Group at North Carolina State University, Raleigh, NC, and joined the graduate program in the Department of Materials Science and Engineering in January 2003. During his graduate studies, he conducted research in the field of silicon wafer defect characterization under the guidance of Dr. Rozgonyi.
ACKNOWLEDGEMENTS

I would first like to thank my advisory committee chair, Dr. George Rozgonyi, for giving me the opportunity to work in his group. He has been exceedingly patient on this time intensive project, built from the ground up. Sincere thanks also goes to my co-chair, Dr. Hans Hallen, whose extensive knowledge in near-field optical microscopy was invaluable to the successful completion of this work, and Dr. Gerd Duscher for his suggestions concerning image analysis. I also wish to thank my postdoctoral advisors who were the driving force of this project, Dr. Abdenacceur Karoui, and Dr. Magnus Wagener, who have instilled a sense of patience, persistence, and precision in the manner I go about my work. A special thanks goes to Dr. Viera Vankova for her assistance with measurements and creation of the NSOM probes.

BP Solar was gracious enough to send us wafers for testing, and their contribution is greatly appreciated. Special thanks goes to Larry Dufour, Rick Lamy, and Terry Briggs in the Burlington Laboratories precision instrument shop, whose fabrication knowledge and boundless patience to put up with my imprecise (or none at all!) schematics helped me to become a more knowledgeable builder.

Lastly, I would like to thank my research group members past and present, Drs. Rossen Yankov, Ernest Van Dyk, Ludwik Kordas, Nathan Stoddard, and Jinggang Lu, as well as Renhua Zhang and Wenjun Zhao, and Yong-Kook Park, who have provided me with scientific collaboration as well as moral support.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
</tbody>
</table>

## CHAPTER 1: INTRODUCTION

1.1 Motivation. ................................. 1
1.2 NSOM Technique. ............................ 4
1.3 Outline. .................................. 8
1.4 References. ............................... 9

## CHAPTER 2: NSOM CARRIER DYNAMICS

2.1 Resolution. ................................. 11
2.2 Laser Beam Induced Current. ................ 16
2.3 Near-Field Optical Beam Induced Current. ... 22
2.4 Free Carrier Absorption. .................... 36
2.5 Near-Field Infrared Free Carrier Absorption. 38
2.6 References ................................. 45

## CHAPTER 3: DESIGN AND CONSTRUCTION

3.1 Laser Optics. ............................... 49
3.2 Scanning. .................................. 54
3.3 Signal Detection. ........................... 68
3.4 Data Acquisition ............................ 72
3.5 References ................................. 82

## CHAPTER 4: RESULTS AND DISCUSSION

4.1 Experimental. ............................... 83
4.2 Microscale Intragrain Defect Analysis. .... 84
4.3 Nanoscale Intragrain Defect Analysis  ....... 88
4.4 References ................................. 100
CHAPTER 5: CONCLUSIONS AND FUTURE WORK

5.1 Conclusions. .................................................. 101
5.2 Improvements and Additions. ............................... 101

APPENDICES

Appendix A: 1D Carrier Diffusion Model. ...................... 103
Appendix B: Shear Force Feedback Gain. ......................... 113
LIST OF TABLES

Table 2.1. Plot characteristics of Figure 2.8. . . . . . . . . . . . . . . . . . . . . . . . . . . 34

Table 2.2 Plot characteristics of Figure 2.9. . . . . . . . . . . . . . . . . . . . . . . . . . . 35

Table 4.1 Plot characteristics of line scans “A” and “B” in Figure 4.7a). . 94
# LIST OF FIGURES

## CHAPTER 1: INTRODUCTION

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Block casting method for multicrystalline silicon.</td>
</tr>
<tr>
<td>1.2</td>
<td>Multicrystalline silicon wafer with silicon nitride coating and front side contacts.</td>
</tr>
<tr>
<td>1.3</td>
<td>Basic NSOM setup.</td>
</tr>
<tr>
<td>1.4</td>
<td>Six common NSOM configurations.</td>
</tr>
</tbody>
</table>

## CHAPTER 2: NSOM CARRIER DYNAMICS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Schematic of light probe formation using a lens.</td>
</tr>
<tr>
<td>2.2</td>
<td>Schematic of light probe formation without a lens.</td>
</tr>
<tr>
<td>2.3</td>
<td>Plane wave propagation across field disturbance U(x,y,0).</td>
</tr>
<tr>
<td>2.4</td>
<td>Charge separation across p-n junction.</td>
</tr>
<tr>
<td>2.5</td>
<td>Geometry of the LBIC model calculation.</td>
</tr>
<tr>
<td>2.6</td>
<td>Schematic of NOBIC operation.</td>
</tr>
<tr>
<td>2.7</td>
<td>Excess carrier density ( n(x) ) at tip position ( x_o ).</td>
</tr>
<tr>
<td>2.8</td>
<td>Aperture size dependence of NOBIC photocurrent.</td>
</tr>
<tr>
<td>2.9</td>
<td>Diffusion length dependence of NOBIC photocurrent.</td>
</tr>
<tr>
<td>2.10</td>
<td>Schematic of ( \tau )-NSOM configuration.</td>
</tr>
<tr>
<td>2.11</td>
<td>( \tau )-NSOM carrier density change due to recombination of defect with size ( 2R ).</td>
</tr>
<tr>
<td>2.12</td>
<td>Comparison of generated photocurrent from NOBIC and ( \tau )-NSOM.</td>
</tr>
</tbody>
</table>
# CHAPTER 3: DESIGN AND CONSTRUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>NSOM system schematic</td>
<td>47</td>
</tr>
<tr>
<td>3.2</td>
<td>Optical bench and protective covering for NSOM system</td>
<td>48</td>
</tr>
<tr>
<td>3.3</td>
<td>Laser optics configuration</td>
<td>49</td>
</tr>
<tr>
<td>3.4</td>
<td>Infrared laser power supply and circuit diagram</td>
<td>50</td>
</tr>
<tr>
<td>3.5</td>
<td>Fiber coupling equipment</td>
<td>52</td>
</tr>
<tr>
<td>3.6</td>
<td>Nomarski images of fiber probe tip</td>
<td>54</td>
</tr>
<tr>
<td>3.7</td>
<td>NSOM scanning system</td>
<td>55</td>
</tr>
<tr>
<td>3.8</td>
<td>Detail of Nanoscope</td>
<td>56</td>
</tr>
<tr>
<td>3.9</td>
<td>X-Y scanning control</td>
<td>61</td>
</tr>
<tr>
<td>3.10</td>
<td>Electrically coupled tuning fork assembly</td>
<td>64</td>
</tr>
<tr>
<td>3.11</td>
<td>Resonance peak shifting during NSOM scanning</td>
<td>65</td>
</tr>
<tr>
<td>3.12</td>
<td>Basic closed-loop feedback block diagram</td>
<td>65</td>
</tr>
<tr>
<td>3.13</td>
<td>Expansion of open loop gain G.</td>
<td>66</td>
</tr>
<tr>
<td>3.14</td>
<td>Circuit board layout and implemented current preamp design</td>
<td>67</td>
</tr>
<tr>
<td>3.15</td>
<td>NOBIC photocurrent collection design</td>
<td>69</td>
</tr>
<tr>
<td>3.16</td>
<td>τ-NSOM signal collection design</td>
<td>70</td>
</tr>
<tr>
<td>3.17</td>
<td>NSOM signal conversion system</td>
<td>72</td>
</tr>
<tr>
<td>3.18</td>
<td>DataTaker front panel</td>
<td>74</td>
</tr>
<tr>
<td>3.19</td>
<td>DataTaker input device front panel</td>
<td>75</td>
</tr>
<tr>
<td>3.20</td>
<td>DataTaker output device front panel</td>
<td>76</td>
</tr>
<tr>
<td>3.21</td>
<td>Resonance Scan front panel</td>
<td>78</td>
</tr>
<tr>
<td>3.22</td>
<td>Scan.vi setup menu</td>
<td>79</td>
</tr>
<tr>
<td>3.23</td>
<td>ImageReader front panel</td>
<td>80</td>
</tr>
</tbody>
</table>
CHAPTER 4: RESULTS AND DISCUSSION

4.1 Minority carrier lifetime map of SiN passivated BP Solar multicrystalline silicon wafer, 125 mm x 125 mm. .......... 85
4.2 Nomarski images of Secco etched sections of wafer in Figure 4.1. .......... 86
4.3 Nomarksi images of the BP mesa structure. ................. 87
4.4 EBIC images of the BP mesa showing electrical activity of intragrain defects. ................. 88
4.5 Images of the TI Ge mesa array used for NSOM calibration. .......... 89
4.6 Images of the TI Ge mesa height variation. ................. 91
4.7 NOBIC images showing linear region of reduced signal. .......... 94
4.8 Correlation of NOBIC contrast to surface topography. .......... 96
4.9 Height measurements of line scans “A” and “B”. ................. 97
4.10 Maps of second line defect. ................... 98
CHAPTER 1  INTRODUCTION

1.1 Motivation

More than 95% of all commercial solar cells are made from silicon. Cast multicrystalline silicon is a major production type of solar grade silicon, accounting for over 40 percent of the total market share in 2003. In the block casting technique, shown schematically in Figure 1.1, the silicon is melted in a separate crucible and then poured into the crystallization crucible, generating blocks with cross sections of more than 60 x 60 cm\(^2\) and weights of over 300 kg. The ingots are crystallized by directional solidification, in which the cooling of crucible walls creates a temperature gradient, where the direction of solid crystal formation is perpendicular to the solid/liquid interface. The heat extraction must occur in a very controlled manner to maintain a high quality of the crystal. In order to achieve a low dislocation density, the melt interface must be kept planar to ensure low thermal stresses. The crystallization and cooling process takes about 30 to 40 hours [1].

![Figure 1.1 Block casting method for multicrystalline silicon. From [1].](image)
The presence of intragrain dislocations due to thermal stress is a well-known phenomenon in cast multicrystalline silicon [2,3]. Such dislocations form from the lateral heat flux of the furnace sidewalls, which control the planarity of the liquid-solid interface, and the vertical heat flux, which controls the solidification front velocity. Convex transverse thermal gradients, although favorable during growth to prevent sidewall nucleation, generate high dislocation densities during cool down near the sidewall edges. Such dislocations either pile up along grain boundaries, or form shallow grain boundaries, thereby acting as recombination centers and degrading the minority carrier lifetime.

In addition to these dislocation clusters acting as recombination centers in bulk defects, they also act as gettering sites for elemental impurities. Impurities are incorporated through the contact of the melt and the crystal with the crucible walls. Typical crucible materials are SiO$_2$ (quartz), Si$_3$N$_4$ as a coating layer on the inside and graphite from the support die. This process introduces oxygen, nitrogen, and carbon in various concentrations. Remnant impurities in the crucible such as the metals Fe, Cr, or Ni also diffuse into melt and solid. The distribution of impurities and doping elements along the growth direction by segregation and the formation of precipitates depend on the growth process and determine the final quality of the crystal [1].

To improve cell efficiencies, solar cell manufacturers require an effective means to minimize the effects of surface and bulk defect recombination centers. One of the most important manufacturing steps of crystalline silicon solar cells is the cost-effective
fabrication of an anti-reflection (AR) coating. Ideally, this coating should not only reduce optical losses but simultaneously provide a reasonable degree of surface passivation and, in the case of multicrystalline silicon, a hydrogen passivation of bulk defects and/or grain boundaries. In recent years it has become increasingly clear that the most promising candidate for the achievement of these vastly differing tasks is silicon nitride generated at low temperature (250 – 450˚C) by means of plasma enhanced chemical vapor deposition (PECVD) [4]. These plasma silicon nitride films act as a surface (and bulk) passivating AR coating, a unique combination that is not met by the standard coatings on crystalline silicon solar cells [5].

Figure 1.2. Multicrystalline silicon wafer with silicon nitride coating and front side contacts.

Previous studies of SiN passivation efficiency in cast multicrystalline silicon have focused on large angle bulk defects, i.e. grain boundaries and dislocations using electron
beam induced current [6], laser beam induced current, [7] or photoconductance decay techniques [8,9]. However, SiN passivation efficiency is not well understood for dislocation clusters, largely because the size of the defect (< 100 nm) is smaller than the spatial resolution of the given techniques. It is therefore desirable to complement existing characterization methods with such that can offer submicron resolution.

Near-field scanning optical microscopy (NSOM) is a submicron resolution characterization technique that provides both qualitative and quantitative information about the carrier lifetime and collection efficiency at small angle grain boundaries. In the NSOM technique, a fiber optic probe with a subwavelength aperture is placed several nanometers from the sample surface, in such a way generating “near-field” light of high spatial resolution. In this work, NSOM is utilized to establish a characterization tool with the ability to map the recombination behavior and carrier lifetime of dislocation clusters in multicrystalline silicon. Two NSOM modes are established to accomplish these goals: 1) laser beam induced p/n junction photocurrent (NOBIC) and 2) minority carrier lifetime based on free carrier absorption of modulated infrared light, (τ-NSOM).

1.2 NSOM Technique

1.2.1 Historical Development

E. H. Synge first developed the NSOM setup as a series of thought experiments in 1928 [10,11]. He identified four potential problems to overcome for the technique to be
successful: 1) generation of an illumination source of sufficient intensity, 2) samples of sufficient surface planarity, 3) regulation of the probe-sample distance, and 4) aperture construction of sufficient diameter. With the invention of the laser, and improvements in sample preparation (e.g. chemical mechanical polishing), the first two of Synge’s concerns were addressed. However, the latter two concerns took the better portion of the century to be sufficiently handled.

![Diagram of NSOM Setup](image)

Figure 1.3 : Basic NSOM Setup. The two major criteria for near field imaging are a subwavelength aperture and a sample-aperture (i.e."gap") distance less than the wavelength of the radiation incident. From [17].

Ash and Nichols [12] in 1972 first used microwave radiation to resolve images with $\lambda/60$ resolution. However, the ability to expand such work to optical wavelengths was hindered by the difficulty of fabricating and controlling subwavelength optical apertures. It was not until the creation of the scanning tunneling microscope (STM) in 1982 by Binning and Rohrer [13] that all of Synge’s concerns were fully addressed. The work of both Ash
and Nichols and Binning and Rohrer used subwavelength apertures created on a flat opaque screen, such as those shown in Figure 1.3. While effective in generating subwavelength radiation, the planarity of the flat screen aperture required very smooth sample surfaces, in the process losing resolution and the potential for topographical data.

A different approach involves subwavelength light sources. These were made by coating pulled pipette tips with metals, and are used for sending light through the apertures at the end (illumination mode) or collecting light from a small region of the sample (collection mode). Since the transmission efficiency of these early tips was small, the resolution was compromised in order to obtain sufficient signal. Betzig and Trautman [14] were able to overcome the deficiencies of the early tips by pulling and coating single mode optical fibers with aperture sizes of < 100 nm, suitable for wavelengths 400 nm - 1500 nm.

A reliable feedback method based on detection of the force interaction between the dithering NSOM tip and sample surface was introduced [15,16]. Using this distance regulation, a topographic image similar to that obtained by a conventional scanning force microscopy (SFM) is acquired simultaneously with the optical image. This provides a way to correlate structural and physical properties at the same sample positions and greatly simplifies interpretation of the NSOM data [17].
1.2.2 Elements of Conventional NSOM

An NSOM system is similar to any other scanning probe microscope. It primarily consists of a tip, in this case a sub-wavelength light source or detector, and a scanning system to move the tip with respect to the sample with sub-nanometer precision. A feedback mechanism is employed to keep tip-sample separation constant and in the near-field regime. A computer controls and synchronizes the data acquisition to the scanning. For NSOM, an external laser and a sensitive optical detection system, which are chosen depending on the specific applications, are added to complete the system. Typically the NSOM head is incorporated in a conventional far-field optical microscope. This allows the positioning of the NSOM tip on the part of the sample of interest. NSOM, therefore, acts as the highest magnification of the entire optical microscope system.

Figure 1.4. Six common NSOM configurations. (a) collection; (b) illumination; (c) collection/illumination; (d) oblique collection; (e) oblique illumination; and (f) dark field. From [18].
As in conventional microscopes, NSOM can be operated in transmission mode, in which the light source and detector are on opposite sides of the sample, and in reflection mode in which the two are on the same side of the sample. Figure 1.4 shows a number of transmission and reflection mode configurations. Using a tapered optical fiber as the NSOM tip, when operating in illumination mode, laser light is coupled into the cleaved, untapered end of the fiber and emerges out of the sub-wavelength aperture at the end of the tapered tip. In collection mode, light is coupled into the fiber through the sub-wavelength aperture and must convert to the propagating mode inside the fiber to reach the detector. In reflection the probe can be used simultaneously as an illumination source and a collector [17].

1.3 Outline

In this work, details concerning the theoretical fundamentals of near-field imaging, instrument construction, and image analysis are presented. Chapter 2 deals with the minority carrier interactions associated with near-field photocurrent and time resolved imaging measurements. Chapter 3 describes the main elements of the NSOM setup, in terms of laser optics, scanning, signal detection, and data acquisition. Chapter 4 describes the results of the microscope in the context of a typical workflow using established characterization techniques (i.e. µPCD, EBIC, Nomarski microscopy). Lastly, Chapter 5 summarizes the findings and describes future work and improvements to the current system.
1.4 REFERENCES


CHAPTER 2: NSOM CARRIER DYNAMICS

In this chapter, the charge carrier processes that generate contrast and resolution in \( \tau \)-NSOM and NOBIC are discussed. First, a general discussion of resolution in near-field and far-field optics is presented. Next, the recombination behavior that determines resolution and contrast in the far-field laser beam induced current and the near-field optical beam induced current methods are described. Finally, an analogous discussion is presented for the far-field free carrier absorption and the near-field infrared free carrier absorption NSOM methods.

2.1 Resolution

Imaging using NSOM is, essentially, a scanning measurement that uses the properties of the near-field light as it interacts with a given material to generate image contrast. To stimulate the discussion of the contrast mechanisms used in this work, a comparison of resolution in far field versus near field optical microscopy is described.

2.1.1 Resolution in Far-Field Optics

Resolution in conventional optics can be described using a combination of classical diffraction theory and Fourier analysis. The main mathematical concepts employed to determine the resolution limit are briefly explained, based on the approach described in [1]. Consider the simplified optical system shown in Figure 2.1 consisting of an aperture of
diameter D, focusing lens L, and image plane P, such that the focal length \( F = LP \) is much longer than the wavelength of the incident plane wave radiation.

Assuming Fraunhofer diffraction from a circular aperture, and in the absence of lens aberration, the intensity profile \( I(r) \) of the focused beam at the image plane is given by the Airy disc equation:

\[
I(r) = \left( \frac{2J_1(\kappa n r \sin \theta)}{\kappa n r \sin \theta} \right)^2
\]

where \( k \) is the wavenumber \( 2\pi/\lambda \), \( n \) is the index of refraction of the medium between the lens and the sample plane, \( r \) is the position of the beam at the sample plane in radial coordinates, \( \theta \) is the half angle of the converging light at the sample plane, and \( J_1 \) is the first
order Bessel function. The first minimum of this equation gives the traditional spatial resolution limit, known as the Rayleigh criterion:

\[ n \sin \theta = (1.22\lambda)/D \]  

(2.2)

From the Rayleigh criterion, spatial resolution is wavelength dependent, and can only be improved by either reducing the wavelength \( \lambda \), or by increasing the numerical aperture \( n \sin \theta \).

![Figure 2.2. Schematic of light probe formation without lens. The illumination of plane wave light of wavelength \( \lambda \) through an aperture of diameter \( D \) is focused onto a sample plane \( P \) located at a distance \( d \ll \lambda \). (Cf. Figure 1.1). From [20].](image)

### 2.1.2 Resolution in Near-Field Optics

To compare the far-field case, consider now the optical system shown in Figure 2.2. The sample plane \( P \) is now at a distance less than the wavelength of the incident plane wave radiation. To simplify the calculation, we can consider the illumination to satisfy the Kirchoff diffraction boundary conditions, namely that the plane wave field distribution is undisturbed
across the aperture, and is otherwise zero across the plane containing the aperture. As a result of the boundary conditions, the spatial resolution is wavelength-independent, limited only by aperture size. Thus, for samples located at distances less than the wavelength of the incident radiation, the smaller the aperture size, the higher the resolvable spatial frequency.

![Figure 2.3. Plane wave propagation across field disturbance U(x,y,0).](image)

According to classical diffraction analysis, a near-field threshold should exist between low and high spatial frequency light, which can be determined using Fourier optical analysis [2]. Consider the Cartesian coordinate system in Figure 2.3. Given a plane wave $\lambda$ propagating in the $z$ direction, impinging on the $xy$ plane, the field distribution $U(x,y,0)$ at the plane, is given by the Fourier expansion and its inverse transform:
\[ A_j(0) = \iint U(x,y,0) \exp\left[-j2\pi(f_x x + f_y y)\right] dxdy \]  

(2.3)

\[ U(x,y,0) = \iint A_j(0) \exp\left[j2\pi(f_x x + f_y y)\right] df_x df_y \]  

(2.4)

where \( f_x \) and \( f_y \) are the spatial frequencies of the wave in the x and y directions, and \( A_j(0) = A(f_x, f_y, 0) \) is the complex amplitude of the plane wave.

At a point \( P_1 \) some distance \( z \) from the plane, the field distribution \( U(x,y,z) \) is similarly given by:

\[ U(x,y,z) = \iint A_j(z) \exp\left[j2\pi(f_x x + f_y y)\right] df_x df_y \]  

(2.5)

where \( A_j(z) = A(f_x, f_y, z) \). Since wave propagation must satisfy the Helmholtz equation:

\[ (\nabla^2 + k^2)U = 0 \]  

(2.6)

where \( k \) is the wavenumber, \( 2\pi / \lambda \), the complex amplitude \( A(z) \) can be determined using by solving the differential equation generated from substituting (2.5) into (2.6):

\[ \frac{\partial^2}{\partial z^2} A_j(z) + k^2 \left[1 - f_x^2 - f_y^2\right] A_j(z) = 0 \]  

(2.7)

\[ A_j(z) = A_j(0) \exp\left(jkz\sqrt{1 - f_x^2 - f_y^2}\right) \]  

(2.8)

When the quantity \( (f_x^2 + f_y^2) < 1 \), i.e. low spatial frequencies, Equation (2.8) is of the form:

\[ A_j(z) = A_j(0) \exp(-jkz\beta) \]  

(2.9)

where \( \beta = \sqrt{1 - f_x^2 - f_y^2} \). The radical is a positive quantity, which results in an increasing exponential term. Substituting (2.9) into (2.5), the resulting plane wave \( U(x,y,z) \) is thus a propagating wave for low spatial frequencies:
\[ U(x,y,z) = \iint A_j(0) \exp(-j k z \beta) \exp\left[j 2\pi (f_x x + f_y y)\right] df_x df_y \]  

(2.10)

When the quantity \((f_x^2 + f_y^2) > 1\), i.e. high spatial frequencies, Equation (2.8) is of the form:

\[ A_j(z) = A_j(0) \exp(j k z \beta) \]  

(2.11)

The radical \(\beta\) is now a negative quantity, which results in a decreasing exponential term in (2.11). Substituting (2.11) into (2.5), the resulting plane wave \(U(x,y,z)\) is an \textit{evanescent} wave for high spatial frequencies:

\[ U(x,y,z) = \iint A_j(0) \exp(j k z \beta) \exp\left[j 2\pi (f_x x + f_y y)\right] df_x df_y \]  

(2.12)

Light from evanescent waves is thus the detected illumination generated from near field optics; it is only detectable at a distance plane (<100 nm) very close to the plane of illumination.

### 2.2 Laser Beam Induced Current

The laser beam induced current (LBIC) technique measures defect recombination strength across a p-n junction device, i.e. the solar cell, using the variation of photoinduced current. By employing the photovoltaic effect, thus closely resembling actual solar cell operation, it is a useful method for characterizing solar cell efficiency. A schematic of LBIC operation is given in Figure 2.4.

In LBIC, a laser generates electron-hole pairs across a p-n junction. Close to the junction, free electrons liberated by the photons on the p side of the junction are swept across to the n side and collected by frontside metallic contacts, while the excess holes on the n side
flow to the backside metallic contact on the p side. When the front and backside contacts are connected, this charge separation creates a measurable current across the junction. When the laser encounters a bulk defect, e.g. dislocations, grain boundaries, etc., recombination at the defect generates fewer available charge carriers, and the collected current across the junction decreases.

A discussion of LBIC resolution and contrast, based on the approach of Marek [3], is now presented in order to motivate the discussion of the near-field analogue of this technique, the near-field optical beam induced current (NOBIC) method. As shown in Figure 2.5, the laser is focused on the surface at a distance x from the defect, in Marek’s case, the grain boundary. The intensity profile is assumed to have a Gaussian shape with a characteristic width $\sigma_0$. The light enters the material and widens by the beam divergence in the optically denser medium. The light intensity is attenuated exponentially by $\exp(-\alpha z)$, $\alpha$ being the absorption coefficient for the laser wavelength. The carrier transport is determined
by the value of the diffusion length, L. The grain boundary is a plane of recombination centers with a spatially invariable recombination velocity $\gamma$. The thickness of the diffused surface layer and of the space-charge region of the junction are summed up to give the value $w$. All carriers generated inside the zone $w$ are assumed to be fully collected.

![Geometry of the LBIC model calculation. From [3].](image)

The induced current for a given generation volume is expressed as:

$$I(x) = I_o - I^*(x)$$  \hspace{1cm} (2.13)

where $I^*(x)$ is the current fraction due to the presence of the grain boundary and $I_o$ is the background current, i.e., current in absence of any defect (or current at a large distance from the grain boundary). The contribution $dI^*$ of a grain boundary areal element $dy \, dz$ is:

$$dI^* = \gamma e^{-z/L} q_p(r,z) \, dy \, dz$$  \hspace{1cm} (2.14)
where \( q \) is the electronic charge and \( p \) is the concentration of minority carriers, assuming, an n-type material. The total beam-induced-current is

\[
I^* = \gamma \int_{-\infty}^{\infty} dy \int_{0}^{\infty} dz \, e^{-(z/L)} \, q p(\sqrt{x^2 + y^2}, z)
\]  
(2.15)

Assuming a Gaussian shape for the lateral dependence of the generation volume:

\[
g(r, z') = \frac{A(z')}{2\pi \sigma^2} e^{-\frac{r^2}{2\sigma^2(z')}}
\]  
(2.16)

The function \( g \) represents the generation strength at a given coordinate, the value \( \sigma \) is the width of the Gaussian function at a depth \( z' \), the depth dependence is described by \( A(z') \).

With the form of \( g(r,z') \) as specified by Equation (2.16), the minority carrier density is

\[
p(r, z) = \frac{1}{4\pi D} \int_{0}^{\infty} d\lambda \, \frac{\lambda}{\mu} J_{\phi}(\lambda r) \int_{0}^{\infty} dz' A(z') \times \left( e^{-\Lambda z'} - e^{-\mu z'} \right) e^{\left( -\lambda^2 \sigma^2(z') / 2 \right)}
\]  
(2.17)

Equation (2.17) is deduced from the results from [4] where:

\[
\Lambda = \frac{1}{L}
\]  
(2.18a)

\[
\mu = \sqrt{\lambda^2 + \Lambda^2}
\]  
(2.18b)

The expression for \( p(r,z) \) can further be evaluated using the relation 6.773.3 given by Gradsteyn and Ryzhik [5] and calculating the integral over \( z \):

\[
I^*(x) = \frac{\gamma q}{\pi D} \int_{0}^{\infty} d\lambda \, \frac{\cos(\lambda x)}{\lambda^2} \int_{0}^{\infty} dz' A(z') \times \left( e^{-\Lambda z'} - e^{-\mu z'} \right) e^{\left( -\lambda^2 \sigma^2(z') / 2 \right)}
\]  
(2.19)

The induced current portion \( I^* \) due to the grain boundary is normalized by the background current \( I_0 \) according to:
\[ i^* = \frac{I^*(x)}{I_o} \] (2.20)

The background current, i.e. the current at a large distance from the grain boundary, is given by:

\[ I_o = \int_0^\infty dz' qA(z')e^{-(z'/L)} \] (2.21)

In the case of light excitation, the generation volume has depth dependence:

\[ A(z') = \alpha e^{-\alpha z'} \] (2.22)

where \( \alpha \) represents the absorption coefficient of the semiconductor at the laser wavelength.

The width \( \sigma \) of the Gaussian function is expressed by:

\[ \sigma^2(z') = \sigma_o^2 + \beta^2 z'^2 \] (2.23)

where \( \sigma_o \) represents the diameter at the surface. The generation volume widens due to the divergence of the beam, and the value of \( \beta \) describes the divergence of the beam. With the generation volume according to Eq. (2.22), the LBIC signal is calculated as:

\[ I^*(x) = \frac{\gamma}{\pi D} \int_0^\infty d\lambda \cos(\lambda x) e^{-\lambda^2 \sigma_o^2 / 2} \alpha \int_0^\infty dz' \times e^{-(\mu z')} \left( e^{-\Lambda z'} - e^{\Lambda z'} \right) e^{-\lambda^2 \beta^2 z'^2 / 2} \] (2.24)

We use the integral 3.322.2 in tables by Gradsteyn and Ryzhik [4] and obtain:

\[ I^*(x) = \frac{\gamma a q}{d\sqrt{2\pi}} \int_0^\infty d\lambda \frac{\cos(\lambda x)}{\beta \lambda^3} e^{-\lambda^2 \sigma_o^2 / 2} \times \left( \exp\Phi_1 \text{erfc}\Phi_1 - \exp\Phi_2 \text{erfc}\Phi_2 \right) \] (2.25)

where:

\[ \Phi_1 = \frac{\alpha + \lambda}{\sqrt{2} \beta \lambda} \] (2.26a)
\[ \Phi_2 = \frac{\alpha + \mu}{\sqrt{2} \beta \lambda} \]  

(2.26b)

are introduced. In case of light excitation the background current \( I_o \) results from Equations (2.21) and (2.22) to:

\[ I_o = \frac{\alpha q}{\alpha + \Lambda} \]  

(2.27)

Finally, we want to evaluate the influence of the space-charge region and the surface layer. The carriers excited within the space-charge region are swept away in the electric field; the probability of recombination at the grain boundary is negligible. The diffused \( n^+ \) layer at the surface is thin compared to the diffusion length and exhibits a short lifetime. The collection efficiency is assumed to be unity in this region. The current \( I_o^\prime \) collected in this region of the thickness \( w \) contributes only to the background \( I_o \) and not to the LBIC signal \( I^* \). The collected \( I_o^\prime \) is represented by:

\[ I_o^\prime = \int_0^w dz' qA(z')e^{-(z'/L)} = \frac{\alpha q}{\alpha + \Lambda}(e^{w(\alpha + \Lambda)} - 1) \]  

(2.28)

due to the fact that the recombination is neglected in this region. We obtain the new background current as:

\[ I_o = \frac{\alpha q}{\alpha + \Lambda}e^{w(\alpha + \Lambda)} \]  

(2.29)

The addition of the zone \( w \) implies that the diffusion problem has the edge of the space-charge region as the upper-boundary. The value of \( \sigma_o \) as needed for the evaluation of Equation (2.25) is the width of the beam at this position. The width \( \sigma_o \) is therefore larger than the beam diameter \( \sigma_o \) as measured on the surface of the sample.
2.3 Near-Field Optical Beam Induced Current (NOBIC)

The near-field optical beam induced current (NOBIC) technique [6-10] is the near-field equivalent of the LBIC method, shown schematically in Figure 2.6. A subwavelength aperture placed 10-20 nm from the silicon wafer is used to generate the near-field light used in the measurement. Modifications to the Marek model, based on the work of Xu, et al. [11] are described to establish the theoretical framework of NOBIC resolution and contrast.

![Figure 2.6. Schematic of NOBIC operation.](image)

The NOBIC image is a photocurrent map generated from optically excited charge carriers across a p-n junction due to near field illumination. With NOBIC, the increase in resolution compared to LBIC is achieved due to the use of the evanescent waves that compose the near field light. The steady-state continuity equation associated with the photoexcited carrier density generated from the near-field beam is:
\[
\frac{\partial n}{\partial t} = \frac{D}{e} \nabla \cdot \mathbf{J} - \frac{n}{\tau} + S = 0 \tag{2.30}
\]

where \( n \) is the photoexcited carrier density, \( D \) is the diffusion coefficient for silicon, \( e \) is the electron charge, \( \tau \) is the recombination lifetime, \( S \) is the excess carrier generation rate and \( \mathbf{J} \) is the current density. With respect to Figure 2.6, the sample surface is represented by x-y plane, and the collected photocurrent is measured in the z-direction.

The photoexcited carrier density is encompassed within an excitation volume, which diffuses outward within the material as the carriers inside the volume return to their relaxed state or recombine. If, as a first approximation, we model the excitation volume as a spherical volume of photons generated from the fiber optic tip, then those photons that do not penetrate the sample will not participate in carrier excitation, which results a hemispherical volume within the material. Although the penetration depth for \( \lambda = 632 \) nm light is approximately \( 1/\alpha = 3.5 \, \mu\text{m} \), the effective interaction depth \( \Delta z \) is on the order of tens of nanometers, which represents the depth of the \( n^+ \) and depletion regions. This reduction of the z direction effectively makes the steady-state equation (2.30) a two-dimensional problem in the x-y plane.

The steady-state equation is again reduced, to one dimension, given the linear nature of the intragrain defect. Although there is a net current flux under the tip in directions along the defect, it will not vary significantly with probe tip movement in the x-y plane. It will resemble the leakage of carriers across the buried interface and quantitatively appear in the
time constant. When the probe tip moves towards or away from the defect, however, the carrier profile will vary strongly in form and magnitude. It is this one-dimensional problem that is considered in this work.

The electron current density $J$ consists of contributions from drift and diffusion currents:

$$
\bar{J} = e \mu \bar{E} n + e D \nabla n
$$

(2.31)

where $\mu$ is the electron mobility. Since there is no $E$ field in the x-y plane, the current density in this plane is due to diffusion current. In the z direction, however, the $E$ field of the p-n junction makes the drift current the dominant contribution in this direction. Xu, et al., used the two-dimensional carrier diffusion model to determine the photocurrent behavior. That is, by using the diffusion current generated from the radial distance between the tip of the probe and the center of the defect $r$, they solved for the excess carrier density, $n(r)$, and integrated the current density $J$ over the x-y plane, to determine photocurrent $N(r)$ in the z direction at each point on the x-y plane. Formulaically, this is represented as:

$$
\nabla^2 n(r) - \frac{n(r)}{D \tau(r)} + \frac{S(r)}{D} = 0
$$

(2.32)

$$
I = \int \int \bar{J} dx dy = \int \int d^2 r = \int n_n(r) e \psi_{dr} d^2 r = \frac{N_n(r)}{w} e \frac{w}{\tau_{dr}} = \frac{N_n e}{\tau_{dr}}
$$

(2.33)

In this work, however, the problem is reduced to a one-dimensional problem. The intragrain defects under investigation are linear defects; the sample is moved with respect to
the tip to produce a series of photocurrent line scans perpendicular to these defects. Equations (2.32) and (2.33) thus simplify to:

$$\frac{\partial^2 n(x)}{\partial x^2} - \frac{n(x)}{D\tau(x)} + \frac{S(x)}{D} = 0$$  \hspace{1cm} (2.34)

$$I = \int \bar{J} dx = \int \frac{n(x)}{w} e\nu_{dr} dx = \frac{N(x)}{w} e \frac{w}{\tau_{dr}} = \frac{N(x)e}{\tau_{dr}}$$  \hspace{1cm} (2.35)

where, \( n(x), \tau(x) \) and \( S(x) = S_o \delta(x-x_o) \) are the carrier density, carrier lifetime due to drift current across the p-n junction, and electron generation rate per unit length, respectively.

To solve for the carrier density \( n(x) \) in (2.34), the carrier lifetime \( \tau(x) \) must be defined. Both diffusion and drift current contribute to carrier lifetime. For this work, the drift lifetime, which is on the order of hundreds of picoseconds due to the thin depletion region of the p-n junction, is much smaller than the diffusion lifetime, which is on the order of several milliseconds. This makes the drift current lifetime the dominant process contributing to the carrier recombination rate, since the rate is proportional to the inverse of the lifetime.

Using this information, the carrier lifetime behavior of the defect is modeled as a step function, that is:

$$\tau(x) = \begin{cases} 0, & 0 < x < a \\ \tau_{dr}, & a < x < \infty \end{cases}$$  \hspace{1cm} (2.36)

where \( a \) is one-half of the total defect size, and \( \tau_{dr} \) is the drift current lifetime. Equation (2.34) then becomes a non-homogeneous second order differential equation of the form:
The solution to equation (2.34), using the lifetime condition of equation (2.36), is shown schematically in Figure 2.7. It is arrived at by adding the homogeneous and particular solutions of equation (2.34), as evaluated in regions I, II, and III. If the carrier density distribution \( n(x) \) shown in Figure 2.7 is described as a piecewise function:

\[
 n(x) = \begin{cases} 
 0, & 0 < x < a \\
 \exp(x - x_o)/D\tau, & a < x < x_o \\
 \exp(-x + x_o)/\sqrt{D\tau}, & x_o < x < \infty 
\end{cases}
\]  

(2.38)

then the homogeneous solution is given as:

\[
 n_h(x_I) = A\exp(x - x_o)/\sqrt{D\tau} + B\exp(-x + x_o)/\sqrt{D\tau}, \quad 0 < x < a \\
 n_h(x_{II}) = C\exp(x - x_o)/\sqrt{D\tau} + E\exp(-x + x_o)/\sqrt{D\tau}, \quad a < x < x_o \\
 n_h(x_{III}) = C\exp(x - x_o)/\sqrt{D\tau} + E\exp(-x + x_o)/\sqrt{D\tau}, \quad x_o < x < \infty
\]  

(2.39)

and the particular solution is given as:
\[ n_p(x) = I \int_{x_o}^{x} f(x) \, dx \]
\[ n_p(x_i) = 0; \quad 0 < x < a \quad (2.40) \]
\[ n_p(x_p) = l \exp(x - x_o/\sqrt{D}t); \quad a < x < x_o \]
\[ n_p(x_{il}) = l \exp(-x + x_o/\sqrt{D}t); \quad x_o < x < \infty \]

where A, B, C, E, and I are solution variables to be determined.

These solutions must satisfy the boundary conditions of the function. In doing so, the unknown variables in the homogeneous and particular solutions are defined. Using the boundary conditions:

\[ n(x), \ n'(x) \text{ continuous at } x = a, \text{ and } x = x_o; \quad n(x) = 0, \quad 0 < x < a, \text{ and } n(\infty) = 0 \]

\[ A = 0 \]
\[ B = 0 \]
\[ C = 0 \]
\[ E = -l \exp(2(a - x_o)/\sqrt{D}t) \quad (2.41) \]

are determined based on the continuity of \( n(x) \). Using the derivative continuity \( n'(x) \), the constant I is calculated as:

\[ \frac{S(x)\sqrt{Dt}}{2D} = I \quad (2.42) \]

Adding the homogeneous and particular solutions, and substituting in the values for the variables, results in a Green function solution, which describes the carrier density distribution as it emanates from a point source:
where $x_o$ is the fiber optic tip position. Integrating each $n(x)$ expression over each interval and adding terms gives an excess carrier density $N(x)$ of:

$$N(x) = S(x)\tau \left( 1 - \exp\left( \frac{a - x}{\sqrt{D\tau}} \right) \right) \left( H(x - a) + H(a - x) \right)$$ (2.44)

where $H(x)$ represents the Heaviside step function.

In reality, the fiber optic tip is not a point source. The carrier density $N(x)$ is a photogenerated population from an aperture, therefore it is desirable to correlate the light intensity from the fiber optic probe to the carrier density. The Airy disc equation (Equation 2.1) can be approximated as a Gaussian distribution:

$$G(x_o - x) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left( -\frac{(x_o - x)^2}{2\sigma^2} \right)$$ (2.45)

where $\sigma$ approximately represents the radius of the fiber optic tip aperture. To correlate the light intensity to the carrier density, a convolution $C(x)$ of the two expressions is done:

$$C(x) = \int_{-\infty}^{\infty} G(x_o - x)N(x)dx$$ (2.46)

The carrier density $N(x)$ is a two term expression due to the step function $H(x)$ terms, which breaks up the integral into two parts:
\[ C(x) = \frac{S(x)\tau}{\sqrt{2\pi\sigma^2}} \left[ \int_a^\infty \exp \left( \frac{-(x_o - x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{a - x}{L} \right) \right) \, dx + \int_{-\infty}^a \exp \left( \frac{-(x_o - x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{a - x}{L} \right) \right) \, dx \right] \]

(2.47)

where \( L = \sqrt{D\tau} \). Changing variables in the second integral gives:

\[ C(x) = \frac{S(x)\tau}{\sqrt{2\pi\sigma^2}} \left[ \int_a^\infty \exp \left( \frac{-(x_o - x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{a - x}{L} \right) \right) \, dx + \int_{-\infty}^a \exp \left( \frac{-(x_o + x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{a - x}{L} \right) \right) \, dx \]

(2.48)

so that:

\[ C(x) = \frac{S(x)\tau}{\sqrt{2\pi\sigma^2}} \left[ \int_a^\infty \left\{ \exp \left( \frac{-(x_o - x)^2}{2\sigma^2} \right) + \exp \left( \frac{-(x_o + x)^2}{2\sigma^2} \right) \right\} \left( 1 - \exp \left( \frac{a - x}{L} \right) \right) \, dx \] \]

(2.49)

Multiplying the terms in Equation (2.49) generates four exponential terms in the integral. Two of the terms will be exponentials multiplied by one, two by another exponential term. In order to fully evaluate the terms, we complete the square twice for those terms which are multiplied by exponential terms, i.e. the \( \exp(\ldots) \cdot \exp(\ldots) \) terms in which the values inside the parentheses are added together. This ultimately yields a convolution \( C(x) \) which consists of an error function solution with a scaling factor, encompassed in the first term of Equation (2.52):

\[ C(x) = \frac{S(x)\tau}{2\sqrt{2\pi\sigma^2}} \left[ \text{erfc} \left( \frac{a - x_o}{\sqrt{2\sigma}} \right) + \text{erfc} \left( \frac{a + x_o}{\sqrt{2\sigma}} \right) - \exp(\Phi_1)\text{erfc}(\phi_1) - \exp(\Phi_2)\text{erfc}(\phi_2) \right] \]

(2.50)

where:
\[
\Phi_1 = \left( \frac{a - x_0}{L} + \frac{\sigma^2}{2L^2} \right), \quad \Phi_2 = \left( \frac{a + x_0}{L} + \frac{\sigma^2}{2L^2} \right),
\]
\[
L = x_0 - \left( \frac{\sigma^2}{L} \right), \quad \mu = x_0 + \left( \frac{\sigma^2}{L} \right)
\]

From Equation (2.35), the total photocurrent \( I \) is determined by integrating the convoluted current density \( C(x) \) across the line scan, i.e.

\[
I = \frac{C(x)e}{\tau_{dr}}
\]

As the tip position \( x_0 \) approaches the defect, the decrease in the convoluted integral results in a decrease in the photocurrent. An estimate of the number of excess carriers available for the NOBIC and \( \tau \)-NSOM measurements through the fiber optic aperture is desired. This estimate by La Rosa [12] is based on the solution of the steady state carrier diffusion equation (Equation 2.34). For negligible surface recombination and a pump absorption length small compared to the diffusion length for carriers, the one-dimensional steady state spatial distribution of excess carriers is given by:

\[
n(x) = \frac{\phi_m}{\sqrt{D\tau}} \exp\left( -\frac{x}{\sqrt{D\tau}} \right)
\]

where, \( \phi_m \) is the input flux density in photons/cm\(^2\) sec, \( D \) is the diffusion coefficient for silicon, and \( \tau \) is the carrier lifetime. Following the calculations of La Rosa [12], the input flux density is defined as the photon emission rate from the excitation laser, \( J_{in} \), divided by the square of the aperture size 2R. The rate of photon emission from a 1 mW, He-Ne excitation laser is:
\[ J_m = \frac{E}{h \nu} = \frac{1 \text{ mW}}{(6.63 \times 10^{-34} \text{ J} \cdot \text{s})/(3 \times 10^8 \text{ m/s})/(632.8 \text{ nm})} = 3.2 \times 10^{15} \text{ photons/sec} \] (2.54)

However, this is not the actual photon emission rate; due to the probe aperture size, only a fraction of the total photons reach the sample surface. Assuming a throughput efficiency of \(10^{-3}\) for an aperture of \(2R = 200\) nm, and a reflection coefficient for silicon of 0.3, the emission rate becomes:

\[ J_{\text{NOBIC}} = (3.2 \times 10^{15} \text{ photons/sec})(10^{-3})(1 - 0.3) = 2.2 \times 10^{12} \text{ photons/sec} \] (2.55)

and the input flux density becomes:

\[ \phi_{\text{in}} = J_{\text{NOBIC}}/a^2 = (2.2 \times 10^{12} \text{ photons/sec})/(0.200 \times 10^{-4} \text{ cm})^2 = 0.5 \times 10^{22} \text{ photons/cm}^2 \text{ sec} \] (2.56)

With \(D = 10 \text{ cm}^2/\text{s}, \tau = 1 \text{ ms}, \) and \(\phi_{\text{in}}\) as defined in Equation (2.56), the carrier distribution underneath the probe \(n(0)\) using Equation (2.53) is:

\[ n(0) = 0.5 \times 10^{22} \text{ carriers/cm}^3 \] (2.57)

Since carriers are created in pairs \(n = p\), the total number of carriers is twice the value indicated in Equation (4.14). This one-dimensional analysis is somewhat justified in a case where the laser beam spot is larger than the diffusion length. In NSOM we have a very localized excitation region and diffusion of carriers is more critical. Therefore a three-dimensional analysis is used. For the three-dimensional analysis, the Green function appropriate to Equation (2.34) is:

\[ \nabla^2 G(x-x_o) - \frac{G(x-x_o)}{D\tau(r)} + \delta(x-x_o) = 0 \] (2.58)
Omitting the boundary conditions in order to simplify the problem, the Green function is given by:

\[ G(x - x_o) = \frac{1}{4\pi(x - x_o)} \exp\left(-\frac{x - x_o}{\sqrt{D\tau}}\right) \] (2.59)

Thus, the corresponding solution for Equation (2.36) at \( x = 0 \) is:

\[ n(0) = \int \left[ \frac{1}{4\pi|x|} \exp\left(\frac{-|x|}{\sqrt{D\tau}}\right) \right] S(x) \frac{D^3}{4D\Delta V} \int \left[ \frac{1}{4\pi|x|} \exp\left(\frac{-|x|}{\sqrt{D\tau}}\right) \right] |x| \sin\theta d\theta d\varphi \] (2.60)

where \( \Delta V \) defines a hemispherical carrier generation volume. Switching to spherical coordinates and defining an equivalent hemispherical domain of radius \( R = (\Delta V)^{1/3} \) gives:

\[ n(0) = \frac{J_{in}}{D\Delta V} \int \int \int R^2 \sin\theta d\theta d\varphi \] (2.61)

generating a carrier density of:

\[ n(0) = 1.6E16 \text{ carriers/cm}^3 \] (2.62)

for a generation volume of \( \Delta V = \) (aperture size)(penetration depth) = 1.4 E-13 cm³. La Rosa’s analysis was done for a \( \tau \)-NSOM measurement in which the penetration depth was taken to be \( 1/\alpha = 3.5 \) \( \mu m \); due to the smaller penetration depth of the p-n junction, a smaller NOBIC carrier population is expected.

According to the analysis of Xu, et al. [11], four factors influence image resolution and contrast:
1) The lateral size of the signal reduction around the defect, denoted by \( \Delta \), defined as the full width at half maximum (FWHM) of the line cut,

2) The defect contrast \( C \), defined as the signal difference between the defect-free region, \( I_b \) and the center of the defect, \( I_d \) divided by the signal in the defect-free region, that is:

\[
\frac{I_b - I_d}{I_b} 
\]  

(2.63)

3) The sharpness, which describes how quickly the signal recovers to the defect-free value. It is defined as the ratio between \( \Delta \) and the distance from the defect where the signals recover 90% of the contrast. The closer this ratio is to one, the sharper the feature is.

4) The total signal, which is determined by the measured photocurrent as calculated by equations (2.36) and (2.37).

The two-dimensional based results of Xu are now compared to the developed one-dimensional, line scan model for similarities and differences with respect to the measurement variables. With respect to these factors, Xu concluded that:

1) \( \Delta \) is determined by the larger of the tip aperture size \( 2R \) or the defect size \( 2a \). The contrast \( C \) increases when \( R \) decreases. However, a decrease of \( R \) causes the
**background signal to decrease.** Figure 2.8 shows the photocurrent plots, generated using Equation (2.54), as a function of decreasing aperture size, $2R$. The linescan model shows that the contrast increases as aperture size decreases, which ultimately increases the lateral resolution $\Delta$. A summary of the plot characteristics is given in Table 2.1.

![Figure 2.8. Aperture size dependence of NOBIC photocurrent.](image)

<table>
<thead>
<tr>
<th>$2R$ (nm)</th>
<th>$I_b$ (nA)</th>
<th>$I_d$ (nA)</th>
<th>C(%)</th>
<th>HM (nm)</th>
<th>$x_{HM}$ (nm)</th>
<th>$\Delta = 2x_{HM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>632.8 ($\lambda$)</td>
<td>58.77</td>
<td>51.53</td>
<td><strong>12.32</strong></td>
<td>55.15</td>
<td>381.0</td>
<td>762.0</td>
</tr>
<tr>
<td>316.4 ($\lambda/2$)</td>
<td>58.77</td>
<td>45.13</td>
<td><strong>23.21</strong></td>
<td>51.95</td>
<td>201.5</td>
<td>403.0</td>
</tr>
<tr>
<td>158.2 ($\lambda/4$)</td>
<td>58.77</td>
<td>35.42</td>
<td><strong>39.73</strong></td>
<td>47.10</td>
<td>115.0</td>
<td>230.0</td>
</tr>
<tr>
<td>79.1 ($\lambda/8$)</td>
<td>58.77</td>
<td>24.28</td>
<td><strong>58.69</strong></td>
<td>41.53</td>
<td>73.50</td>
<td>147.0</td>
</tr>
</tbody>
</table>
For the intragrain defects analyzed in this study, a reasonable approximation for the defect size is on the order of a few lattice spacings. For silicon, the lattice spacing, \( a_{\text{Si}} = 0.5431 \) nm, so the defect size, \( 2a = 20 a_{\text{Si}} \sim 10 \) nm is used as a rough estimate. Due to limitations in the fiber optic probe construction, the smallest attainable tip aperture size ranges from about 50-100 nm. Thus, the tip aperture size \( 2R \) will always be larger than the defect size, and is the dominant factor in determining image resolution.

![Diagram showing photocurrent vs tip position for four different diffusion lengths](image)

Figure 2.9. Diffusion length dependence of NOBIC photocurrent.

Table 2.2 Plot characteristics of Figure 2.9. As diffusion length \( L \) increases, the sharpness ratio \( \Delta / x_{10C} \) decreases, indicating the photocurrent profile becomes less sharp. Parameters used for the plots are defect size, \( 2a = 10 \) nm, diffusion coefficient, \( D = 10 \) cm\(^2\)/s, aperture size, \( 2R = 158.2 \) nm (\( \lambda/4 \)), and photon generation rate, \( S = 4.064 \times 10^{11} \) / nm sec.

<table>
<thead>
<tr>
<th>( L ) (nm)</th>
<th>( \tau_{\text{dr}} ) (ps)</th>
<th>( I_b ) (nA)</th>
<th>( I_d ) (nA)</th>
<th>( C ) (%)</th>
<th>( I_{\text{HM}} ) (nA)</th>
<th>( x_{\text{HM}} ) (nm)</th>
<th>( \Delta ) (nm)</th>
<th>( I_{10C} ) (%)</th>
<th>( x_{10C} ) (nm)</th>
<th>Sharp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.5</td>
<td>58.77</td>
<td>35.42</td>
<td>39.73</td>
<td>51.95</td>
<td>115.2</td>
<td>230.4</td>
<td>56.43</td>
<td>220</td>
<td>1.04</td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>58.77</td>
<td>24.30</td>
<td>58.65</td>
<td>47.10</td>
<td>147.0</td>
<td>294.0</td>
<td>55.32</td>
<td>315</td>
<td>0.93</td>
</tr>
<tr>
<td>200</td>
<td>40</td>
<td>58.68</td>
<td>14.76</td>
<td>74.85</td>
<td>41.53</td>
<td>211.2</td>
<td>422.4</td>
<td>54.28</td>
<td>530</td>
<td>0.79</td>
</tr>
<tr>
<td>316</td>
<td>100</td>
<td>57.78</td>
<td>10.11</td>
<td>82.50</td>
<td>33.95</td>
<td>282.5</td>
<td><strong>565.0</strong></td>
<td>53.01</td>
<td>744</td>
<td>0.75</td>
</tr>
</tbody>
</table>
2) The diffusion length in the defect-free region $l_0$ determines the sharpness of the defect. The greater $l_0$ is, the longer it takes for the signal to recover to the defect-free value. Figure 2.9 shows the photocurrent plots, generated as a function of increasing background diffusion length. The linescan model verifies Xu’s results; as diffusion length increases, the plots become broader, resulting in decreased sharpness. A summary of the plot characteristics is given in Table 2.2.

3) The diffusion length of the defect $l_1$ strongly affects contrast, but not $\Delta$ or sharpness. The line scan model uses a step function (Equation (2.38)) to model the carrier lifetime. Because of this approximation, the lifetime at the defect is zero, and consequently the diffusion length of the defect $l_1$ is zero. This is not physically feasible for an actual defect, and further modeling analysis must be done to account for the defect lifetime.

4) Changing the diffusion constant $D$ is equivalent to changing $l_1$ and $l_0$ at the same rate. This result is confirmed in Figure 2.9. Although the figure uses a constant diffusion coefficient while varying the carrier lifetime, varying the diffusion coefficient while keeping the lifetime constant has the same effect on diffusion length $l_0$.

2.4 Free Carrier Absorption

The free carrier absorption lifetime method is a truly a noncontacting technique, relying on optical electron-hole pair (ehp) generation and on optical detection using two different wavelengths. A pump beam using photons with energy $h\nu > E_g$ creates ehps. The
signal, as read by a photodectector, is based on the dependence of the free carrier absorption of photons with $h\nu < E_g$ on the density of free carriers. The probe beam transmitted photon flux density $\Phi_t$ is given by:

$$\Phi_t = \Phi (1 - R)^2 \frac{\exp(-\alpha_{fc}d)}{1 - R^2 \exp(-2\alpha_{fc}d)}$$

(2.64)

where $\alpha_{fc}$ is the free carrier absorption coefficient, $d$ the sample thickness, and $R$ the reflectance. For n-type semiconductors $\alpha_{fc}$ is given as:

$$\alpha_{fc} = K_n \frac{\lambda}{n}$$

(2.65)

Where $K_n$ is a materials constant and $\lambda$ the wavelength of the probe beam. For n-Si, $K_n \sim 10^{-18}$ cm$^2$/μm$^2$ and for p-Si $K_p \sim (2-2.7) \times 10^{-18}$ cm$^2$/μm$^2$.

The method can be used in both steady state and transient modes. In the steady-state mode, a probe beam is incident on the sample, while the transmitted beam is detected by an infrared detector. The pump beam is chopped at a few hundred Hz for synchronous detection by a lock-in amplifier. In the transient mode, the pump beam is pulsed, and the time-dependent carrier density is detected through the transmitted probe beam.

The change in the transmitted probe beam as a result of a chopped or pulsed pump beam is:

$$\Delta \Phi_t = -(1 - R) d \Phi_0 \Delta \alpha_{fc} \frac{1}{1 + R}$$

(2.66)
using \( \exp(-2\alpha_{fc}d) \sim \exp(-\alpha_{fc}d) \sim 1 \) in Equation (2.64) with \( \alpha_{fc}d \ll 1 \). The change in the absorption coefficient is:

\[
\Delta \alpha_{fc} = K_n \lambda^2 \Delta n = \frac{K_n \lambda^2}{d} \int_0^d \Delta n(x) \, dx
\]

The change in carrier density, \( \Delta n \), is an important parameter, since it relates to the minority carrier lifetime, surface recombination velocity, sample reflectivity, pump beam absorption coefficient, and the photon flux density. The transient version data interpretation is simpler since the transient carrier decay, monitored by the transient absorption coefficient, contains the recombination information. The probe and pump beams can be perpendicular to one another and, by scanning the probe beam, it is possible to map the lifetime through the wafer thickness [13].

### 2.5 Near-Field Infrared Free Carrier Absorption (\( \tau \)-NSOM)

Imaging silicon with photons from infrared (IR) light, whose energy is below the bandgap energy of silicon \( (E_g = 1.12 \text{ eV}) \), can be compared to imaging window glass with visible light, since in both cases the materials are transparent to the incident radiation. The infrared image, however, can be altered through the introduction of photoexcited carriers created by visible light of photon energy larger than the silicon bandgap energy. Although the fate of carriers photoexcited into bands of extended states is complicated by a number of effects such as surface recombination and diffusion, a sufficient number of charge carriers are available for free carrier absorption. That is, photoexcited electrons thermalize to states in the bottom of the conduction band, where they may absorb infrared photons. This absorption
pumps the carriers into higher energy states in the conduction band, and consequently, the IR transmission of the sample falls.

A number of competing processes serve to deplete the population of photoexcited carriers. Depending on the sample examined, these processes can have characteristic times from picoseconds to milliseconds. Under continuous wavelength (cw) bandgap illumination, equilibrium is quickly reached and the population of photoexcited carriers under any given position will not change. Under pulsed illumination, however, the population of these carriers will change as a function of time and location. The IR transmission will likewise change over similar time scales and distances. Through monitoring of the IR transmission, the behavior of photoexcited carriers can therefore be measured. By changing the frequency at which photoexcited carriers are created, various transient mechanisms of the sample can be probed. Insofar as this behavior is a function of the defect population of silicon, changes in IR transmission can thus be used to monitor semiconductor quality.

LaRosa, et al. [14] developed an NSOM system, shown in Figure 2.10, to extend the spatial resolution to subwavelength dimensions. Different optical responses of the sample were probed by performing cw and ac-modulated experiments. By tuning the ac component, they showed that a number of sample response mechanisms can be studied. In the experiment outlined in [14], pulsed visible light ($\lambda = 0.632 \, \mu m$, 1 mW power) is used to generate free carriers, and continuous wavelength infrared light ($\lambda = 1.15\mu m$, 1 mW power) is used to create an image of the surface of a polished, oxidized silicon wafer, with an average lifetime
of 1.6 ms. At high visible light modulation frequencies (~20 kHz), insufficient relaxation
time is provided between pulses to allow for relaxation of the free carriers and the change in
the transmitted IR signal is small. However, at low chopping frequencies (~1 KHz), the
system has considerable time to relax in each cycle, the infrared signal change increases. This
transmitted IR is detected synchronously with a lock-in amplifier, and at each pixel, the
signal from a number of pulses is averaged [15].

With NSOM resolution, one must treat carefully the concept of any local property of
excess electrons, such as lifetime, due to the high mobility of the electrons and holes. With a
diffusion constant $D = 10 \text{cm}^2/\text{s}$ and $\tau > 1 \mu\text{s}$, the carriers can travel as much as $\sqrt{Dt} > 30$
$\mu\text{m}$. Therefore, the lifetime may seem to be a quantity intrinsically “averaged” over a
relatively large area of a sample. However, using the analysis of La Rosa [14], a localized
lifetime, which can characterize defects with a submicron resolution, is established.

Figure 2.10. Schematic of $\tau$-NSOM configuration. From [14].
The qualitative picture, as shown in Figure 2.11, is that the excess carriers that are locally created by the visible light rapidly diffuse into a region larger than the size of an image and form a distribution $n$, weakly peaked under the probe. This distribution reflects the local lifetime through its variations $\Delta n$, which are sensed at high spatial resolution by the IR. Although one might first guess that $\Delta n$ would vary on a length scale given by the diffusion length, it is not the case. Rather, $\Delta n$, and hence the resolution, varies on a length scale given by $R$, the size of the “object” or inhomogeneity in the sample.

![Figure 2.11. τ-NSOM carrier density change due to recombination of defect with size 2R.](image)

As an example, let us consider some domain of radius $R$ within which the carriers recombine by $\delta(1/\tau)$ faster than elsewhere. A faster decay will lead to a variation in the local carrier density, $\Delta n$. This gradient causes an influx of carriers $\sim D \cdot R \cdot \Delta n$ from the surrounding volume. In steady state, the decay balances the influx, and the concentration deviates from a uniform value by:

$$\frac{\Delta n}{n} = -\delta(1/\tau) \left(\frac{R^3}{3D}\right) r^{-1}$$  \hspace{1cm} (2.68)
at a distance \( r > R \) from the center; the function is localized for small \( R \) (small defect areas), which are of interest here. In the defect-rich area \( \Delta n/ n \sim R^2 \ d(t^{-1})/3D \), which produces a detectable change in IR absorption. The change in IR signal is inversely proportional to the mobility and decreases for smaller feature sizes, thus passing the problem of high resolution to the one of signal level because the small \( \Delta n' \)’s have to be distinguished [14].

Since free carrier absorption in \( \tau \)-NSOM is proportional to the excess carrier density, we compare the \( \tau \)-NSOM carrier density generated under the tip to that generated due to NOBIC photocurrent. To determine the carrier density for \( \tau \)-NSOM, we again use the solution for the steady-state carrier diffusion model in Equation (2.34):

\[
\begin{align*}
    n(x_I) &= 0, & 0 < x < a \\
    n(x_{II}) &= \frac{S(x)\sqrt{D\tau}}{2D} \left( \exp\left( x - x_o / \sqrt{D\tau} \right) - \exp\left( 2a - x - x_o / \sqrt{D\tau} \right) \right); & a < x < x_o \\
    n(x_{III}) &= \frac{S(x)\sqrt{D\tau}}{2D} \left( \exp\left( -x + x_o / \sqrt{D\tau} \right) - \exp\left( 2a - x - x_o / \sqrt{D\tau} \right) \right); & x_o < x < \infty
\end{align*}
\]

However, the region of interest is the generation volume due to tip illumination. In this case, \( x = x_o \) in Equation (2.43) so that the solution becomes:

\[
\begin{align*}
    n(x_I) &= 0, & 0 < x < a \\
    n(x_{II}) &= \frac{S(x)\sqrt{D\tau}}{2D} \left( 1 - \exp\left( 2(a - x_o) / \sqrt{D\tau} \right) \right); & a < x < x_o \\
    n(x_{III}) &= \frac{S(x)\sqrt{D\tau}}{2D} \left( 1 - \exp\left( 2(a - x_o) / \sqrt{D\tau} \right) \right); & x_o < x < \infty
\end{align*}
\]

and the total carrier density \( N(x) \) is:
\[
N(x) = S(x)\tau \left(1 - \exp\left(\frac{2(a - x_o)}{\sqrt{D\tau}}\right)\right)\left[H(x - a) + H(a - x)\right] \tag{2.70}
\]

Following the convolution procedure described in Section 2.3, the convoluted carrier density is given as:

\[
C(x) = \frac{S(x)\tau}{2\sqrt{2\pi}\sigma^2} \left[\text{erfc}\left(\frac{a - x_o}{\sqrt{2}\sigma}\right) + \text{erfc}\left(\frac{a + x_o}{\sqrt{2}\sigma}\right) - \exp(\Phi_1)\text{erfc}(\phi_1) - \exp(\Phi_2)\text{erfc}(\phi_2)\right] \tag{2.71}
\]

where:

\[
\Phi_1 = \left(\frac{2(a - x_o)}{L} + \frac{\sigma^2}{2L^2}\right), \quad \phi_1 = \left(\frac{a - \zeta}{\sqrt{2}\sigma}\right), \quad \zeta = x_o - \left(\frac{2\sigma^2}{L}\right) \tag{2.72a}
\]

\[
\Phi_2 = \left(\frac{2(a + x_o)}{L} + \frac{\sigma^2}{2L^2}\right), \quad \phi_2 = \left(\frac{a + \mu}{\sqrt{2}\sigma}\right), \quad \mu = x_o + \left(\frac{2\sigma^2}{L}\right) \tag{2.72b}
\]

Figure 2.12. Comparison of current generated from NOBIC (blue) and τ-NSOM (red). The τ-NSOM contrast is scaled in green to match the NOBIC contrast. Parameters used for the plots are defect size, \(2a = 10\) nm, diffusion length, \(L = 100\) nm, diffusion coefficient, \(D = 10\) cm\(^2\)/s, aperture size, \(2R = 158.2\) nm (\(\lambda/4\)), lifetime, \(\tau_{dr} = 10\) ps, photon generation rate, \(S = 4.064\times10^{11} / \text{cm}^3\text{ sec.}\)
A comparison of the convoluted current densities associated with NOBIC and τ-NSOM, as calculated in Equations (2.47) and (2.71), is shown in Figure 2.12. In addition, the τ-NSOM contrast is scaled to match the NOBIC contrast, shown in green in Figure 2.12. This is accomplished by subtracting the difference of the NOBIC and τ-NSOM current values at \( x = 0 \) from the τ-NSOM profile, then multiplying the resulting values by \( \sqrt{2} = 1.41 \). The contrast difference in the NOBIC and unscaled τ-NSOM profiles is due to two factors. In the τ-NSOM convolution in Equation (2.71), both the first term of \( \Phi \) and the second terms of \( \mu \) and \( \zeta \) shown in Equations (2.72a) and (2.72b) vary from the terms in NOBIC convolution in Equation (2.49) by a factor of two. The signal change due to the defect is encompassed in the terms.

In summary, a description of near-field light, in terms of its evanescent components was given. These evanescent waves are used to generate a photoinduced current across a p-n junction, in this case a solar cell, in the NOBIC mode, and to measure a frequency-dependent carrier lifetime due to infrared free carrier absorption in the τ-NSOM mode. Since the fiber optic tip can only generate a percentage of the total photons available for imaging, a convolution of carrier density and Gaussian light intensity is done to correlate tip illumination to the number of expected carriers available for NOBIC photocurrent or τ-NSOM free carrier absorption. These models will be used as the basis of the data analysis in Chapter 4.
2.6 REFERENCES


CHAPTER 3: DESIGN AND CONSTRUCTION

This chapter describes the construction of the near-field microscope for measurement in τ-NSOM and NOBIC modes. The NSOM system is divided into four major subsystems: 1) laser optics, which involves the alignment and generation of near field light, 2) scanning, which controls the three dimensional movement of the fiber optic tip, 3) signal detection, which collects and amplifies the signal and, 4) data acquisition, which converts the analog signals to digital data. The entire NSOM system is shown schematically in Figure 3.1.

Figure 3.1. NSOM system schematic.
The nanoscope and laser optics sit on top of an optical bench, shown in Figure 3.2, which is isolated from low frequency building vibrations by means of six individually pressurized inner tubes. In addition, the laser optics and nanoscope are housed in an enclosure to prevent excess light, dust, and wind from negatively influencing the measurement. Other NSOM equipment (e.g. Nanoscope control unit, power supplies, etc.) is stored underneath the table on specially designed shelves.

Figure 3.2. Optical bench and protective covering for NSOM system.

In τ-NSOM mode, the carrier excitation beam (HeNe, 632 nm) and infrared (1310 nm) transmission beams are coupled into a single fiber optic probe of subwavelength aperture. Visible light is absorbed by the sample in the process of carrier generation, while
the transmitted infrared intensity through the sample is collected by an objective and measured using a photodetector. In NOBIC mode, the junction current generated from optically excited electron hole pairs is measured directly and amplified. The procedure for generating the NSOM image is now described, by detailing each of the major subsystems.

![Laser optics configuration](image)

**Figure 3.3.** Laser optics configuration.

### 3.1 Laser Optics

The first step in creating the NSOM image is to generate near field illumination. Laser optics in NSOM includes all associated equipment involved in generating, aligning, and coupling of the light sources into the fiber optic probe used for near-field illumination. The major equipment includes the lasers, positioning equipment, and fiber couplers, as
shown in Figure 3.3. Of greatest importance is the construction of the fiber optic probe, which is discussed in a later section.

### 3.1.1 Lasers

During the construction process, several types of visible lasers were used for the excitation source, including both gas and laser diode lasers. In all cases, the excitation wavelength was 632 nm. In Figure 3.2, a 0.9 mW continuous wavelength (CW) gas HeNe laser was used. In addition, a 632 nm 1 mW HeNe laser from Aerotech, Inc. was used. In both cases however, the laser output was found to be too unstable over long periods of time. In its current configuration, a 632 nm continuous wavelength (CW), 3 mW HeNe laser diode from Meredith Instruments is used.

The infrared laser used was a 1310 nm CW Mitsubishi 725B8F 10 mW AlGaAs laser diode. The infrared diode required the construction of a power supply, since the operating...
characteristics (operating voltage = 1.22 V, operating current = 18 mA) were very precise.

The power supply circuit and its circuit schematic is shown in Figure 3.4.

3.1.2 Positioning

As shown in Figure 3.3, the positioning equipment consists of three multi axis stages (i.e. movement in the X,Y and Z directions), and a rack and pinion slide for precise linear movement, all from Edmund Optics. Two of the stages are used for the alignment of the visible laser, the third is used for the infrared laser. The visible laser is fixed on the bench, defining the optic axis for alignment. One XYZ stage is attached to the optical bench and is used to adjust the focusing lens, the second stage is an XZ stage attached to the rack and pinion slide, which moves in the Y direction. In addition, tilt stages from CVI are attached to the z axis of both stages for fine control. The IR diode is also fixed on the table; the stage for the infrared laser is an XZ stage attached to the table.

3.1.3 Fiber Optic Coupling

Fiber coupling involves both coupling the visible and IR beams into separate fibers and joining them together into the subwavelength probe. This is accomplished using a wavelength division multiplexer (WDM), shown schematically in Figure 3.5a. The WDM used was a 2 x 2 50/50 broadband coupler from Newport Corporation. In this configuration, the visible and IR beams are coupled into two separate fibers and are coupled together into two outgoing fibers. One of the outgoing fibers is coupled into the fiber optic probe using a
Siecor fiber splice, (Figure 3.5b) the other is used to monitor the stability of the laser power output.

![Figure 3.5. Fiber coupling equipment. a) Schematic of broadband coupler (WDM). b) Siecor fiber optic splice.](image)

### 3.1.4 Fiber Optic Probe

The common method of generating near-field illumination involves coupling light into a metal coated, fiber optic probe of subwavelength aperture, as developed by Betzig, et al. [1]. Fiber tips are created by two separate methods: mechanically pulling the fiber while heated using a CO₂ laser, or by chemical etching. Typical fiber tip diameters range from 50 to 200 nm. Since the index of refraction decreases in the tapered region, total internal reflection is lost in an uncoated fiber. To counteract this effect, the tips are coated with thermally evaporated aluminum. Care must be taken in the evaporation process, which involves tilting and rotation of the fibers, to ensure a smooth, uniform aluminum layer. Poorly evaporated films result in uncoated areas of fiber due to the formation of metal
islands, i.e. “pinholes” that act as leakage sites for the light and reduce the light throughput from the aperture.

Even with a properly coated fiber probe, problems still exist. The first obvious disadvantage is the loss of light output through the aperture, generally about a $10^5$ to $10^6$ decrease in light output from the input light intensity. In addition, the aluminum coating absorbs the light near the tip and creates localized tip heating [2]. Experimental measurements showed that the Al coating near the NSOM aperture can reach temperatures as high as 470˚ C in operating conditions [3]. Consequently, the input laser power cannot be arbitrarily increased in order to boost the output power at the NSOM aperture. With optimal coupling, a safe working input power is 0.5 - 1 mW. Significantly above this level, the Al at the end of the tip melts, causing the tip to catastrophically fail [4].

The tips created for this work were chemically etched using both the protection layer method patented by Turner [5] and by the tube etching process [6]. In the protection layer method, the buffer layer is stripped off the fiber, leaving the glass core exposed. The core is then dipped in hydrofluoric acid (HF) covered by an organic cover layer, and systematically pulled from the HF layer. The HF etches away the glass while the meniscus formed by the organic-acid interface creates the taper. Although this method is sufficient for creating tips, care must be taken to avoid vibrations in the fiber, and to maintain a constant pull rate, so as to avoid irregularities in the tip cone angle.
An alternative method used to eliminate the tip irregularities seen in the protection layer method is the tube etching process, in which the buffer layer remains intact and the tip remains submerged in the HF layer. It has been shown [6] that the quality of the cone angle and taper smoothness improves with this method in comparison to the protection layer. Nomarksi images of the tips created by both methods are shown in Figure 3.6.

![Nomarski images of fiber optic tips](image)

Figure 3.6 Nomarski images of fiber optic tips. a) Tip (far left) and detail created using protection layer method. b) Tip created using tube etching method.

### 3.2 Scanning

The following section details the NSOM horizontal and vertical scanning control, as described in the operation manual [7]. Vertical control, by means of shear force feedback, enables the probe tip to be in close contact to the sample surface (~10 nm) in order to
generate the evanescent waves used in imaging. Horizontal control requires a precision piezoelectric element, controlled by a modified scanning tunneling microscope (STM). Figure 3.7 shows the NSOM scanning system, consisting of the nanoscope and the control unit.

![Figure 3.7. NSOM scanning subsystem.](image)

**3.2.1 Nanoscope**

The NSOM system uses a modified scanning tunneling microscope (STM) system from Digital Instruments. The Nanoscope I consists of two major components, the nanoscope and the control unit. The nanoscope, shown in detail in Figure 3.8, consists of the scanning
piezoelectric tube (PZT), which controls the motion of the fiber optic probe, the head which holds the piezo, the base which supports the head and sample, the support structure for the base, coarse adjust screws, the stepper motor to control the fine adjust screw, and the tuning fork preamplifier. The control unit contains the low voltage power supply, the high voltage power supplies, the feedback board, the X-Y board, the stepper motor board, and the front panel.

Figure 3.8. Detail of Nanoscope.

The function of the instrument is to scan the tip in an X-Y raster fashion across the sample, maintaining a constant distance from the sample. A feedback loop uses a setpoint
current as a reference input to control the height of the tip by applying a voltage to the scanning piezo (the Z drive voltage). A shear force feedback loop, based on the voltage output of an oscillating tuning fork, is used to control the height of the tip. The operator has control over the setpoint current, but does not have direct control over the height of the tip above the sample – that is adjusted by the feedback, but is typically less than 10 angstroms.

A. Scanning Head

The scanning head controls the 3-dimensional motion of the tunneling tip. The removable head consists of an Invar shell into which is mounted the piezo tube scanner, which controls the tip. In order to minimize vertical thermal drifts, it is important to have a good thermal match between the piezo crystal and its mount. To reduce this thermal drift, the shell is made of Invar, a nickel steel which has an expansion coefficient of 0.04 microns/degree over a length of 3/4”, close to that of the expansion coefficient of the piezo crystal. The combination of the Invar and piezo have a net expansion coefficient so that fairly large temperature changes can be tolerated without the vertical piezo drive going out of range. The shell is nickel and chrome plated. The head is held in position with magnets. Pressed into the top of the Invar shell is a ceramic ring into which is epoxied the 1/2” diameter piezo crystal.

The nanoscope uses a 1/2” diameter single tube piezo crystal to control the three-dimensional motion of the tip. The tube piezo provides both rigidity and simplicity and is the design of choice for tunneling microscopes. The piezo tube has six separate electrodes
driven by five separate drivers. This design gives X and Y motions which are perpendicular, provides good sensitivity, minimizes horizontal and vertical coupling, and mounts the tip where it is easy to visualize.

The vertical motion of the crystal is controlled by the inner electrode, called the Z electrode, which is driven by the feedback loop. The X and Y scanning motions are controlled by two electrodes each, the pair being driven with opposite polarities, i.e. there are two drivers for X and two for Y. These electrodes are called –Y, -X, Y, X starting at the front of the crystal and going clockwise as seen from the top. With respect to Figure 3.8, the Y electrode is the white wire on the piezo which faces the reader. A positive voltage on X (and therefore a negative voltage on –X) will cause the tip to move to the right, as seen from the top of the scanning head. Similarly, a positive voltage on Y will cause the tip to move downwards, i.e. towards “6 o’clock”, as seen from the top of the head.

Piezoelectric materials convert electrical energy to mechanical energy. The PZT uses the electrical signals sent from the control unit of the nanoscope to generate mechanical strain, moving the tube in the x, y, and z directions. The piezoelectric characteristics vary for different types of materials and tube sizes, so it is necessary to determine the range in which a given tube will move to determine the maximum scan size and vertical range. The important factor for these calculations is the strain or d constant, which relates the amount of mechanical strain to the amount of applied voltage. For this application, the direction of the applied voltage is perpendicular to the strain movement, so the appropriate d constant is $d_{31}$,
where the 1, 2 and 3 directions correspond to the x, y, and z axes. The nanoscope PZT is an EBL #4 from Stavely Incorporated, which has a $d_{31}$ of $-97 \times 10^{-12} \text{m/V}$, where the negative sign on the $d$ constant denotes that the PZT contracts when a positive voltage is applied to it. The maximum scan limits are thus determined using the following relations for x, y, and z range:

\[ \Delta x = \frac{0.90d_{31}V_x L^2}{D t} = (19 \text{ nm/V})(300 \text{ V}) = 5.7 \mu m \]  
(3.1a)

\[ \Delta y = \frac{0.90d_{31}V_y L^2}{D t} = (19 \text{ nm/V})(300 \text{ V}) = 5.7 \mu m \]  
(3.1b)

\[ \Delta z = \frac{d_{31}V_z L}{t} = (5.5 \text{ nm/V})(300 \text{ V}) = 1.6 \mu m \]  
(3.1c)

where:

- $\Delta x, \Delta y, \Delta z$ = PZT movement range
- $V_x, V_y, V_z$ = applied voltage range = $\pm 150 \text{ V} = 300 \text{ V}$
- $d_{31}$ = strain constant = $-97 \times 10^{-12} \text{m/V}$
- $L$ = PZT tube length, 2 in
- $D$ = PZT tube diameter, 0.5 in
- $t$ = PZT wall thickness, 0.35 in

The NSOM system is thus limited to $5.7 \times 5.7 \mu m^2$ scans with a maximum height variation of $1.6 \mu m$. The scan size and height calibration of the NSOM system will be verified in further detail in Chapter 4.

**B. Base and Base Support**

The head is supported by an Invar base into which are mounted three 1/4-80 precision screws for the course and fine adjustment of the tip height. Invar is used to match the
expansion coefficient of the head. The screws are positioned in a triangular arrangement in which the forward screws have a mechanical advantage of about one for raising or lowering the tip, and the rear screw, which is used by the stepper motor for fine adjustment, has a mechanical advantage of −0.12. That is, raising the screw lowers the tip. The screws have magnetized balls in their ends to hold the head down securely.

**C. Stepper Motor**

The rear screw supporting the head is driven by a stepper motor. It rotates at 800 steps per revolution, and advances the screw by 0.4 microns per step. With the mechanical advantage of 0.12, the tip advances at about 500 Angstroms per step. The hex drive from the stepper motor to the screw has a backlash of about ten steps, so when single stepping the tip in one direction and then reversing, it will take about 10 steps for the tip to reverse its direction. The exact number of steps is measured by observing the Z drive voltage while single stepping the tip. Furthermore, the stepper motor can be used to calibrate the Z-axis of the piezo in Angstroms per volt, by single stepping the tip an even number of steps, and recording the change in the Z drive voltage. The voltage change is divided by the total number of steps to get the volts per step.

**D. X-Y Scanning Circuitry**

The X and Y axes are driven with triangular waveforms generated on the X-Y BOARD on the bottom shelf of the control unit. The portion of the front panel is shown in Figure 3.9. The X and Y scan circuits are identical except for the frequency of their
waveforms. The range of the X drive frequency is nominally from 3-150 Hz and Y from .02 to 1 Hz. The actual ranges are somewhat larger, but the frequency will always be indicated properly on the front panel digital meter. The nanoscope is typically run with X at 50 Hz and Y at .05 to .1 Hz.

Figure 3.9. Nanoscope X-Y scanning control. The scan voltage, frequency, amplitude, and offset are adjusted from the front panel.

The scan circuitry uses standard function generator chips that produce 10 V triangular waveforms whose amplitude is adjusted with the AMPLITUDE knobs. To these waveforms are added offsets that can be read on the digital panel meters. The offset knobs are wired so that turning them clockwise sends a lower offset. This is done so that a clockwise rotation of the X offset moves the image to the right. However, to move the image to the right, the tip
must be moved to the left by applying a lower offset voltage. Similarly, a clockwise rotation of the Y offset applies a lower offset voltage, moving the tip downwards and the image up.

The amplitude of the scan is controlled by a 10-turn knob which has a readout on its face. The amplitude of the triangle waveform has a range of ±10 V on the low scale, or a total peak-to-peak range of 20 V. This number does not display well on a 10-turn knob, so the display is given as one half of scan voltage displayed on the knob, with a maximum reading of 1000. The reading on the knob must therefore be multiplied by .02 to give the total scan voltage. The range switch is marked x2 in the low range position. A 1 volt scan would be a setting of 050 on the AMPLITUDE knob. On the high range, where a x15 high voltage amplifier is switched into the circuit, the knob reading must be multiplied by 0.60.

The range switch on the scanning circuitry affects both the X and Y axes and puts a gain of 15 on both the offset and scan voltages to the piezo. The range switch does not affect the offset readings on the digital meters, so the readings must be multiplied by 15. The GAIN switch and the AMPLITUDE knobs do not affect the X and Y drive signals to the scope. These signals are 10 volts in amplitude and are affected only by the frequency knobs. In this manner, the image remains constant in size, independent of the scan amplitude. When the scan amplitude is decreased, the image will zoom towards the center with a larger magnification and conversely when the scan amplitude is increased.
In order to insure that the range of the electronics is not exceeded with a combination of large scan and offset voltages, a “clip” circuit and light warns the user when the electronics is or is about to be saturated. When the CLIP light is blinking, the scan amplitude or the absolute value of the offset should be decreased so that the light stops blinking.

On the front panel are external X and Y scanning inputs so that scanning can be controlled by analog outputs from a computer. These inputs must be enabled with the switch on the X-Y BOARD. This switch disables the internal scanning, but the amplitude, offset, and clip circuitry are still functional on the external inputs. To obtain a gain of one from the applied input to the piezo drive, with the gain switch in the x1 position, the amplitude knob should be set to 500. The x1 / x15 switch is still operable then the external inputs are used. The X and Y drives will change when the amplitudes of the external inputs change, changing the size of the image, but will not be affected by the x1/x15 switch. The maximum voltage that can be applied on the external inputs is ± 5V. With the AMPLITUDE knob at 1000 and the gain switch in the x15 position, this gives a maximum voltage to the piezo of ±150 V.

### 3.2.2 Shear Force Feedback

Shear force feedback is the mechanism that controls the tip-sample distance necessary to maintain near-field illumination. The shear force mechanism used in the setup follows the tuning fork method of Karrai and Grober [8], in which a single mode fiber optic probe, as described in section 3.1.4, is attached using adhesive to a piezoelectric tuning fork, which is then attached to a glass slide, as shown in Figure 3.10.
The design of the assembly is such to allow for the greatest amount of tuning fork vibration, while minimizing excessive tuning fork stress. As shown in Figure 3.10b), one tine of the tuning fork is glued to the edge of a glass slide, while the fiber optic probe is glued to the other tine. The contacts of the tuning fork are trimmed or bent prior to gluing so as to facilitate the connection of gold wire between the contacts and conducting wire to the lock-in oscillator.

![Figure 3.10 a) Assembly schematic, with glass slide dimensions. b) Electrically coupled tuning fork assembly.](image)

The resonance frequency of the tuning fork is determined by the shear forces that act upon the optical probe – e.g. the strength of the epoxy from the tuning fork to the probe, atomic force interactions with the probe, etc – which are dependent on the probe-sample distance. As the probe scans across the sample, variations in this distance will register as a shift both in amplitude and frequency of the initial resonance peak. The purpose of the
feedback loop is to adjust the distance of the tip so as to maintain the original peak position, as shown in Figure 3.11.

The shear force feedback loop can be idealized as a basic closed feedback loop, shown in Figure 3.12. In the figure, R equals the input, C represents the output, G represents the open loop gain and E represents the error, i.e. the input and output difference.

\[
C/R = \frac{G}{1+G}
\]  

Figure 3.11. Resonance peak shifting during NSOM scanning.

Figure 3.12. Basic closed-loop feedback block diagram.

The closed loop gain is the ratio of the output to input C/R. With the relationships \( C = GE \) and \( E = R - C \), the closed loop gain in terms of open loop gain, G, is determined as:

\[
\frac{C}{R} = \frac{G}{1+G}
\]  

(3.2)
The question remains as to how to determine the open loop gain. It consists of five stages: 1) a proportional + integral feedback gain from the nanoscope, 2) a high voltage gain controlling the z-axis voltage from the nanoscope, 3) a voltage gain associated with the voltage / height conversion with the piezoelectric tube, 4) the tuning fork preamplifier gain, and 5) the lock-in amplifier gain. This is shown schematically in Figure 3.12, and calculated theoretically in Appendix B.

![Diagram of open loop gain stages](image)

**Figure 3.13 Expansion of open loop gain G.**

Establishing shear force feedback is accomplished in two steps: determining the resonance frequency of the tuning fork, and matching this signal to that of the nanoscope control. In terms of Figure 3.10, the amplitude of the resonance frequency, in millivolts, represents the reference signal, and the current setpoint dial on the nanoscope control unit represents the output.

To determine the resonance frequency, a piezoelectric tuning fork is stepped through a given frequency range (~ 30-45 kHz) driven by an EG&G 7225 lock-in amplifier. The
tuning fork in turn generates an AC voltage, whose output is sent through a preamplifier and read by the lock-in. At the tuning fork resonance frequency, the magnitude of the signal is amplified, resulting in a sharp peak. The amplitude of this peak above the background signal determines the reference voltage necessary for feedback. The output signal from the tuning fork is on the order of 10s of microvolts. For proper detection by the lock in amplifier, the signal must be amplified. An electrically driven current preamplifier was designed for this purpose. The circuit schematic and completed preamp is shown in Figure 3.14.

![Circuit schematic and completed preamp](image)

**Figure 3.14.** a) Tuning fork current preamplifier. b) Implemented circuit design.

The two main factors that determine the quality of a resonance scan are the signal to noise ratio (S/N) and the quality factor, Q, defined as the ratio of the frequency of the maximum amplitude to full width at half maximum (FWHM). The Q factor of free standing tuning forks (i.e. unattached to a probe) such as those in Figure 3.7 are in the thousands, but
with fiber probes attached, the Q factor broadens to around 100 – 500. Other factors affecting S/N and Q are circuit and software related, e.g. varying the oscillator amplitude, adjusting the circuit to different background noise levels, etc. However, these factors can be adjusted with each resonance measurement to improve the quality of the peak.

After determining the resonance frequency for a given tuning fork, the next step is to bring the tip into near field range. In reference to Figure 3.12, this requires the comparison of the lock-in output from the tuning fork R to the setpoint current on the Nanoscope control. The setpoint current controls the applied voltage for vertical movement to the piezoelectric tube, which varies from +150 V (fully retracted with respect to the sample) to −150 V (fully extended). The tip is brought into near field conditions by a combination of the coarse positioning screws and the stepper motor, as described in section 3.2.1. Feedback is established with the tip fully extended (i.e. −150 V), at a “critical point” where the application of more current from the setpoint dial would increase the voltage to +150 V. This corresponds to a point zero error in the block diagram in Figure 3.13. As the tip approaches the sample, the feedback loop pulls back the tip at a point where resonance peak shifting occurs (i.e. non-zero error) so as to maintain zero error as described in Figure 3.12.

**3.3 Signal Detection**

The acquired signal is different in τ-NSOM and NOBIC modes, although the source of the signal in both modes is based on near-field illumination. In τ-NSOM, the signal is a modulated transmitted IR signal, based on free carrier absorption due to near field
illumination of visible light. In NOBIC, the signal is a junction current due to charge separation created by near-field illumination.

### 3.3.1 Signal Collection

Signal detection in NOBIC involves photocurrent collection, amplification and noise filtering. The main challenge in acquiring the NOBIC signal is developing a structure that can collect the charge carriers from both sides of the solar cell p-n junction. The developed design uses silver paint applied to a traditional SEM carbon mount acting as the contact for the p side of the junction, and an insulated wire connected to the n+ side of the junction using gold wire, as shown in Figure 3.15.

![Figure 3.15. NOBIC photocurrent measurement design. The BP Solar mesa structure, as described in Chapter 4, is on the carbon mount.](image)
Signal detection in τ-NSOM involves infrared collection, amplification, and noise filtering. These steps are necessary since the light output is greatly diminished due to the small aperture size of the probe. The design of the nanoscope enables the collection and photodetection of the transmitted infrared light, as shown in Figure 3.16a). A cylindrical copper housing was machined to mount the detector. The housing then sits in the center of a Unitron objective stage, into which the objective is screwed.

![Image of τ-NSOM signal collection design](Image)

**Figure 3.16.** τ-NSOM signal collection design. a) Infrared detection equipment. b) Circuit diagram of InGaAs photodetector. c) InGaAs detector responsivity curve. At the operating frequency, $\lambda = 1310$ nm, the responsivity is approximately 0.800 A/W.
The objective used is a standard achromatic microscope objective from Melles Griot with a resolving power of 63x, numerical aperture of 0.85, working distance of 0.14, and a focal length of 3.1 mm at \( \lambda = 632.8 \) nm. The detector used is a THORLABS FGA10 InGaAs photodetector, shown schematically in Figure 3.16b), which converts the light intensity to a photocurrent according to the formula:

\[
I_p = P_o R(\lambda)
\]

(3.3)

where \( P_o \) is the incident power from the infrared laser diode in watts, and \( R(\lambda) \) is the responsivity of the detector in amperes per watt. At the infrared wavelength of 1310 nm, the responsivity is 0.800 (Figure 3.17b). The output laser power exiting from the aperture is on the nanowatt scale, so nanoamperes of photocurrent are expected for detection.

### 3.3.2 Amplification & Filtering

The incident photocurrent that enters the objective is on the order of tens of nanoamperes. The photocurrent must be amplified in order for the data acquisition card to measure discernable changes in the transmitted infrared intensity (\( \tau \)-NSOM), or junction photocurrent (NOBIC) as it scans the sample. To accomplish this, the photocurrent from the detector is connected to a Keithley 428 current amplifier, which can increase the gain by up to a factor of 1 \( \times \) 10.
3.4 Data Acquisition

Data acquisition for the NSOM system is carried out in two stages: signal conversion, involving both digital to analog and analog to digital conversion, and data analysis using specialized software. Each stage is described in detail.

Figure 3.17. NSOM signal conversion system. It consists of the National Instruments SCB-68 shielded connector block, which connects to the NI PCI-6014 card in the computer.

3.4.1 Digital-Analog Signal Conversion

Signal conversion in the NSOM system consists of two items: a SCB-68 68-pin shielded connector block, and a PCI-6014 data acquisition card, both from National
Instruments, shown in Figure 3.17. The connector block, which plugs into the card at the back of the computer, supplies signal connection to analog inputs and outputs. The data acquisition card both sends digitized signals to the connector block outputs and digitizes the analog signals received from the connector block inputs. In addition, software is installed on the computer to monitor the output signals read in from the connector.

The signals from the analog inputs sent to the connector are the z output from the nanoscope control unit, used in topography mapping, and the IR intensity (in τ-NSOM mode) or photocurrent (in NOBIC mode) measured from the sample as read by the photodetector or current amplifier, respectively. The analog outputs are connected to the X and Y auxiliary inputs on the nanoscope control unit. The scan voltages sent from the data acquisition card to these inputs control the movement of the PZT during scanning.

The connector is designed to measure and send analog signals between −10 V to +10 V. This range is suitable for scan voltages, and for photocurrent/junction current, however, the Z output sent from the nanoscope control unit ranges from −150 V to +150 V, requiring the use of a voltage divider.

3.4.2 Acquisition Software

The acquisition software, DataTaker, was developed by Dr. Hans Hallen in the NCSU Physics Department. It was created using LabView, a graphically based programming language. The front panel is shown in Figure 3.18. DataTaker is a virtual instrument (VI) that
provides both measurement and analysis tools for determining resonance and generating data. An overview the software is now given; however, an acquired NSOM image only requires the use of two VIs from the main front panel: Resonance Scan and Scan. These are described in further detail in separate sections.

DataTaker is a flexible data acquisition tool. It is able to acquire data from a variety of sources in a way that is transparent to the user. The price paid for this flexibility is the necessity for device configuration, which is managed using a device preference file. When DataTaker loads, the software prompts the user for a preference file. Once in the program, clicking the instrument configuration button takes the user to a large panel, shown in Figure 3.19. The instrument configuration is an aid in choosing what to move and by how much, so the conversions don't have to be done by hand.
The input devices are displayed as elements of an array. For each array element, several kinds of entries are used to identify each device, e.g. the name and measured units. The "it is a" entry identifies the type of expected input from a given device. The 'Num' entry defines which array element the 'it is a' entry refers to, so the user can have multiple kinds of the same type of device. The source tells the computer more information about the type of device (e.g. if it is an analog-digital conversion card (ADC), board, lock-in amplifier, etc.). The reading from this device is converted to units by multiplying the reading by 'units/in' then adding 'zero offset.' Other values include whether the device uses a general purpose interface bus (GPIB) or serial port (RS-232) connection, the GPIB address or serial port number, the channel or port, and the expected voltage range, used by the ADC.

The output device window, shown in Figure 3.20, is arranged like the input window. For the NSOM measurement, the outputs sent from the ADC card are the scan voltages for
the X and Y directions of the Nanoscope control unit. In addition to the values mentioned for the input window, there are additional limitations placed on the outputs. These limitations include the max and min values, speed limits in milliseconds, acceleration limits, “Bzzt” step size, and the current value. The speed limit saves the tip when moving to a new point in X and Y, but will slow down the scan if it is set too low. When moving the output value, the user can set it to “Bzzt” (i.e. continuously move) rather than jump to the new value. This is desirable to prevent crashes in scanning X and Y. If the Bzzt step size is too low, the scans will also take longer. A good Bzzt rate is about one Bzzt step per millisecond. Once the input and output devices have been defined, the user can close the instrument configuration window and save the setup as a new preference file.

![Figure 3.20 DataTaker output device panel.](image)

The buttons of the DataTaker front panel bring other useful VIs for running the instrument and analyzing data. Move lets the user change an output, allowing for change before and after the movement, and supporting various stopping conditions. Move and watch
is a more sophisticated version of Move, with more options and the ability to record some inputs. Watch 8 lets the user see the analog inputs as unscaled voltage values. Freq is a mini spectrum analyzer to help determine if there is white noise, and what frequencies are present so as to find the source. It also allows one to grab and save traces from the scope meter, lock-ins, or plug-in card. The lock-in amplifier VIs control the operation of the various lock-in models. Once the user is satisfied with the preliminary analysis using these VIs, NSOM image acquisition can begin [9].

A. Resonance Scan

Resonance Scan, shown in Figure 3.21, is the VI that controls the resonance peak measurement, as described in section 3.2.2. The user controls the initial and final frequency values, the amplitude of the oscillation frequency, the number of and time between measurement points, and the sensitivity and time constant of the lock-in amplifier. When the user presses the “Start” button, the AC voltage, with the user-selected drive amplitude, is stepped through the frequency range, with a frequency step size of determined by the frequency range divided by the number of measurement points.

The voltage vs. frequency information is sent via GPIB cable from the lock-in to the computer, which plots the results in the VI. The voltage and frequency data at each point can thus be determined to find the suitable resonance frequency for NSOM imaging. Furthermore, this data can be saved as an ASCII file, which can be displayed using graphing software (e.g. Microsoft Excel).
B. Scan

The scan.vi is the virtual instrument responsible for acquiring the NSOM image. Clicking on the scan button on DataTaker brings up a scan setup menu, as shown in Figure 3.22. The user first enters a scan name, or does so using the “Autosave” option, which will save the data when the scan finishes using the scan name as a file name. In the setup menu, the scan area is defined by the center position, size, and number of points in the X and Y directions. The final image can be rotated as well, although the movement of the measurement axes cannot. Other options in the setup menu include the choice of what values
(i.e. optical, topography, etc.) to measure at each point, the time between points, and whether to “Bzzt” or jump to each point. Selecting the back data checkbox, generates two images, one in the forward and backward directions. The 'equispace points' option is used to get the same number of DAC steps between each point in a line, or stepper motor steps, etc. Clicking the “Start” button begins the scan, the OK button saves the choices in the preference file, and the cancel button exits the menu.

![Scan.vi setup menu](image)

**Figure 3.22 Scan.vi setup menu.**

**C. ImageReader**

At this point in the measurement, scan.vi has generated one to six images of the τ-NSOM or NOBIC image, surface topography, and laser power output. To view and process the images, the ImageReader VI, shown in Figure 3.23, is used. It is a separate VI from DataTaker, so DataTaker must be stopped before ImageReader is opened. When ImageReader runs, the program prompts the user for a file generated from scan.vi. Once
loaded, the image appears in the left side window, scaled to the units as specified by the scan.vi setup menu.

3.23 ImageReader front panel.

Although a single file is loaded, the number of images generated from the file is noted in the “of this many images” indicator. Since only three inputs are measured during one measurement, an even number of images generally denotes back and forward scans, an odd number means forward scans. For example, if the forward scans of the NOBIC image, topography and laser power were generated from a file, the indicator would read “3”; if back and forward scans were taken, the indicator would read “6”. By clicking on the “look at” selector, and pressing the “Other/Reset Image button”, the user can view each image.
There are various image processing functions that ImageReader can apply to the image. The most important is the “Background Subtract” button. Its purpose is to subtract the difference of the highest and lowest values from an image from the average value so as to “flatten” the image to an average value. It is especially useful for the topographical images, where sample tilt often obscures surface features that can be used to correlate features in the optical images. Aside from the image processing functions, the most important aspect of ImageReader is the line scan feature, which measures the photocurrent/intensity/tip height versus position in both the horizontal and vertical directions. These line scans can be averaged over a given number of lines in the scan. These line scans will form the basis of the analysis in Chapter 4. The theoretical one-dimensional carrier diffusion model established in Chapter 2 will be compared to the actual data from the line scans.
3.5 REFERENCES


CHAPTER 4: RESULTS AND DISCUSSION

The results of this work are divided into two sections: microscale and nanoscale defect analysis. The microscale analysis uses established carrier lifetime characterization techniques to identify areas of high intragrain dislocation densities. The nanoscale analysis extends the characterization to the scale of the intragrain defect clusters.

4.1 Experimental

Carrier lifetime mapping of a 5×5 in² (125 x 125 mm²), SiN passivated cast silicon BP Solar wafer (see Chapter 1, Figure 1.2) was done using an AMECON JANUS 300 microwave photoconductive decay (μPCD) system. A grain boundary map was obtained by optically scanning the sample surface and using image processing software. The grain boundary outline was then overlaid onto the mPCD map to correlate the microstructure with the lifetime map. Regions A and B, shown in Figure 4.1, were chosen because of the presence of both high and low lifetime areas, both within (Region A) and across (Region B) high angle grain boundaries. These regions were cut, polished, Secco etched for 60 seconds to delineate the dislocations, and observed using Nomarski differential interference contrast microscopy [1].

For NSOM analysis, sections from the same silicon block with dislocation clusters were cut and sent to BP Solar for phosphorus annealing to create p-n junctions. To delineate
areas of interest on the sample, p-n “mesas” was created. The samples were heated on a hot plate, and wax placed on the region of interest. The sample was then Secco etched, removing the phosphorus layer except for the wax region. The wax was then removed using acetone, creating a p-n “mesa” several mm in diameter. Areas of electrically active intragrain dislocation clusters were then determined in these mesas using electron beam induced current (EBIC). Using the EBIC maps as a guide, NOBIC characterization was then done.

4.2 Microscale Intragrain Defect Analysis

4.2.1 Microwave Photoconductance Decay (µPCD)

The lifetime and grain boundary map are shown in Figure 4.1. Areas of high lifetime (i.e. green to blue areas greater than 10 µs) denote effective bulk passivation. From the overlaid grain boundary map, areas with large structural features, such as high angle grain boundaries and twin boundaries show improved lifetime (green areas of τ ≥ 10+ µs) with passivation.

The low lifetime areas (i.e., yellow to red areas of τ < 10 µs) of the wafer appear to be limited by a different mechanism. The low lifetime band at the top of the map corresponds to light element impurity diffusion (e.g. oxygen, carbon) from the growth crucible, and no significant increase in intragrain defect density was noted in this region. The other low lifetime areas of the wafer are attributed to intragrain defects. Regions A and B were cut for Nomarski microscopy to identify the intragrain fine structure [1].
4.2.2 Nomarski Microscopy

Figure 4.2 shows the Nomarski results from Regions “A” and “B” of the wafer in Figure 4.1. These regions were chosen to investigate the dependence of carrier lifetime to the presence of intragrain dislocations, both within and across grains. The areas of low lifetime in the μPCD maps when examined under Nomarski show the presence of dislocation clusters, creating arrays leading to the formation of low angle grain boundaries. By contrast, the high lifetime regions show virtually no dislocations.
The presence of intragrain defects is further supported based on Nomarski images of the aforementioned mesa structures, as shown in Figure 4.3. Figure 4.3a) shows a mesa before etching. No dislocations are visible on the surface. Visible are dark regions attributed to areas of diffused phosphorus. In Figure 4.3b), the right edge of another mesa was Secco etched for one minute, and the Nomarski images show bands of dislocation clusters similar to those in Figure 4.2.

Figure 4.2. Nomarski images of Secco etched sections of silicon wafer in Figure 4.1.
4.2.3 Electron Beam Induced Current (EBIC)

To confirm that the dislocation clusters are electrically active, electron beam induced current (EBIC) analysis was done on the aforementioned mesa structure. At 300 K, the major contribution to EBIC contrast comes from high angle grain boundaries, although faint contrast due to intragrain defects is seen in the central grain in Figure 4.3a). Reducing the temperature of the sample to 150 K, thus probing energy levels deeper within the bandgap, increases the EBIC contrast of the intragrain defects in the central grain. However, the limited EBIC resolution renders the areas of dislocation arrays as hazy grey contrast, as shown in Figure 4.3b). It is at this point that NOBIC analysis is required. A 5 μm² area, as marked by the red square in Figure 4.3b, is scanned using NOBIC.
4.3 Nanoscale Intrgrain Defect Analysis

4.3.1 Topography Calibration

As a preface to the NOBIC analysis, the horizontal and vertical ranges of the NSOM system are verified using an array of germanium mesas, spaced 2.5 µm apart on a silicon substrate, from Texas Instruments. The Nomarksi image of the array is shown in Figure 4.5a and the NSOM topographical map is shown in Figure 4.5b.

Figure 4.4. EBIC images of the BP mesa showing electrical activity of intragrain defects. a) room temperature EBIC scan, showing electrical activity at the high angle grain boundaries. b) 150 K scan, showing increased contrast at the intragrain defects.
As calculated in Chapter 3, the maximum NSOM scan size is 5.7 µm x 5.7 µm.

This value is determined by multiplying the calculated amount of piezo movement, in nanometers per volt, by the total applied voltage:

\[
\Delta x = \frac{0.90d_{31}V_y L^2}{D t} = (19 \text{ nm/V}) (300 \text{ V}) = 5.7 \mu m
\]  
(4.1a)

\[
\Delta y = \frac{0.90d_{31}V_y L^2}{D t} = (19 \text{ nm/V})(300 \text{ V}) = 5.7 \mu m
\]  
(4.1b)

The total applied voltage from the Nanoscope control unit is ±10 V, for a total range of 20 V.

The voltage range is controlled using the AMPLITUDE knob, which ranges from 0 to 1000.

This requires the values on the knob to be multiplied by .02 to get the actual applied voltage.
In addition, the voltage gain on the control unit can be increased by 15, so that the total voltage range is \( \pm 150 \text{ V} = 300 \text{ V} \).

For the NSOM scan in Figure 4.5a), the value on the AMPLITUDE knob was set at 500, so that the total voltage range is:

\[
(500)(0.2)(15) = 150 \text{ V} = \pm 75 \text{ V} 
\] (4.2)

Using this value for the applied voltage gives a scan size of:

\[
\Delta x = \frac{0.90 d_{31} V_z L^2}{D_t} = (19 \text{ nm/V})(150 \text{ V}) = 2.85 \mu\text{m} 
\] (4.3a)

\[
\Delta y = \frac{0.90 d_{31} V_y L^2}{D_t} = (19 \text{ nm/V})(150 \text{ V}) = 2.85 \mu\text{m} 
\] (4.3b)

Given that the array consists of 2.5 \( \mu\text{m} \) diameter mesas, spaced 2.5 \( \mu\text{m} \) apart, as seen in Figure 4.5a, it is reasonable to assume that the NSOM image in Figure 4.5b is that of a single mesa, with the edge of a second mesa appearing on the right side of the image.

For the vertical calibration, we again turn to the height calculation in Chapter 3. In this case, we use:

\[
\Delta z = \frac{d_{31} V_z L}{t} = (5.5 \text{ nm/V})(300 \text{ V}) = 1.6 \mu\text{m} 
\] (4.4)

Figure 4.5a shows the NSOM image in Figure 4.5b in color to accentuate the height variations. A three-dimensional rendering of the mesa, based on the topographical pixel values of Figure 4.6a, is shown in Figure 4.6b.
The vertical range of the piezo tube was calculated as $1.6 \, \mu m = \pm 800 \, nm$. A voltage divider must be used in the coaxial cable from the Nanoscope control unit to the data acquisition card, which can only read in voltages ranging from $\pm 10$ volts. The divider in this system was intentionally capped to range from $\pm 133.6 \, V = \pm 8.91 \, V$ so as to minimize hysteresis effects associated with large cyclical vertical piezo movement. The consequence of capping the voltage range is that the height measurement saturates, which accounts for the white region in Figure 4.6a), and the flat light blue region in the three dimensional projection of Figure 4.6b).

Using profilometry, the height of the mesas was measured as 800 nm. With this guideline, it is approximated that the mesa is encompassed by the region surrounded by the green band in Figure 4.6a, and the area above the yellow region of Figure 4.6b. This assertion
is further supported by the fact that the mesa array was Secco etched for 1 minute, so the array substrate that would ordinarily have been present, was etched away, accounting for the areas of “negative height”, with respect to the piezo range, as seen in Figure 4.6b).

4.3.2 Near-Field Optical Beam Induced Current (NOBIC)

The NOBIC analysis extends the microscale defect analysis to the submicron scale. Figure 4.7 shows the scanned NOBIC image, which is composed of 100 line scans in the vertical (i.e. “X”) direction. The measured photocurrent in the z-direction is proportional to the number of photogenerated excess carriers. As described in Chapter 2, the photocurrent $I(x)$ is the integral of the current density distribution evaluated at each point on the line scan, i.e.:

$$ I = \frac{C(x)e}{\tau_{dr}} \quad (4.5) $$

where the convoluted carrier density $C(x)$ is:

$$ C(x) = \frac{S(x)\tau}{2\sqrt{2\pi}\sigma^2} \left[ \text{erfc}\left(\frac{a-x_o}{\sqrt{2}\sigma}\right) + \text{erfc}\left(\frac{a+x_o}{\sqrt{2}\sigma}\right) - \exp(\Phi_1)\text{erfc}(\phi_1) - \exp(\Phi_2)\text{erfc}(\phi_2) \right] \quad (4.6) $$

with:

$$ \Phi_1 = \left(\frac{a-x_o}{L} + \frac{\sigma^2}{2L^2}\right), \quad \phi_1 = \left(\frac{a-\xi}{\sqrt{2}\sigma}\right), \quad \xi = x_o - \left(\frac{\sigma^2}{L}\right) \quad (4.7a) $$

$$ \Phi_2 = \left(\frac{a+x_o}{L} + \frac{\sigma^2}{2L^2}\right), \quad \phi_2 = \left(\frac{a+\mu}{\sqrt{2}\sigma}\right), \quad \mu = x_o + \left(\frac{\sigma^2}{L}\right) \quad (4.7b) $$

Figure 4.7 shows two line scans of the NOBIC image, labeled “A” and “B”, picked for analysis using the photocurrent model. The characteristics of the photocurrent model
versus the actual photocurrent line scan data are given in Table 4.1. The data acquisition card digitizes the 20 volts (i.e. ±10 V range) and assigns them to a discrete value from $\pm 2^{15} = \pm 32768$. The conversion of the raw value read in by the data acquisition card to the measured photo current value is given as:

$$I = \left[ \frac{\text{(raw value)}}{32768} \right] \left( \frac{10 \text{ V}}{10^9 \text{ nA} / \text{A}} \right) \left/ \left( \text{current amplifier gain}, 10^8 \text{ A/V} \right) \right.$$  (4.8)

These particular line scans were chosen to highlight the validity of the photocurrent model, in spite of the drop in background signal from line scan “A” to “B”, due to power fluctuations in the visible excitation laser. Assuming the reduction of photocurrent is due to defect recombination, the “X” position for each line scan is offset so that the defect position at the lowest photocurrent value lies at zero. Consequently, the range of the fiber optic tip movement should be considered as some positive or negative distance away from the defect, as noted in Figure 4.7c), as opposed to a length. That is, the “X” range for a full scale NSOM scan is ±2.85 µm, not 5.7 µm.

Using the analysis from La Rosa [2] from Chapter 2, the number of generated carriers in the NOBIC measurement is estimated. The carrier density under the probe $n(0)$ is:

$$n(0) = \frac{J_{in}}{D\Delta V} \iiint \left[ \frac{1}{4\pi R} \exp \left( \frac{-|R|}{D\tau} \right) \right] R^2 dR \sin \theta d\theta d\varphi$$  (4.9)

where $J_{in}$ is the photon emission rate, laser = 2.2 E 12 photons/sec, $\Delta V$ is the hemispherical carrier generation volume, and $R$ is the radius associated with this volume = $\Delta V^{1/3}$. Using the model aperture size $2R = 390.2$ nm listed in Table 4.1, and the $n^+$ layer thickness, measured to be 300 nm using bevel polishing, the model generation volume is:
Table 4.1. Plot characteristics of line scans “A” and “B” in Figure 4.7a).

<table>
<thead>
<tr>
<th>C(x) values</th>
<th>S (x) (1/nm³ s)</th>
<th>L (nm)</th>
<th>D (cm²/s)</th>
<th>τ_dr (ps)</th>
<th>2R (nm)</th>
<th>2a (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4.62 E 11</td>
<td>31.6</td>
<td>0.1</td>
<td>100</td>
<td>390.2</td>
<td>10</td>
</tr>
<tr>
<td>B</td>
<td>4.41 E 11</td>
<td>31.6</td>
<td>0.1</td>
<td>100</td>
<td>390.2</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plot values</th>
<th>I_b (nA)</th>
<th>I_d (nA)</th>
<th>C (%)</th>
<th>Δ (nm)</th>
<th>x_{10C} (nm)</th>
<th>Sharp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A data</td>
<td>58.75</td>
<td>51.50</td>
<td>12.3</td>
<td>607</td>
<td>514.4</td>
<td>1.18</td>
</tr>
<tr>
<td>A model</td>
<td>58.95</td>
<td>51.52</td>
<td>12.6</td>
<td>480</td>
<td>424.7</td>
<td>1.13</td>
</tr>
<tr>
<td>B data</td>
<td>56.13</td>
<td>49.06</td>
<td>12.6</td>
<td>430</td>
<td>457.4</td>
<td>0.94</td>
</tr>
<tr>
<td>B model</td>
<td>56.30</td>
<td>49.19</td>
<td>12.6</td>
<td>470</td>
<td>431.1</td>
<td>1.09</td>
</tr>
</tbody>
</table>

Figure 4.7. NOBIC images showing linear region of reduced signal. a) Photocurrent maps, 4.8 μm x 4.8 μm. Line scans “A” and “B” are chosen to show the decrease in background signal. b) 2D plots of NOBIC data from line scans and photocurrent model.
Using the values for diffusion length from Table 4.1, \( D = 0.1 \text{ cm}^2/\text{s}, \tau = 100 \text{ ps}, \) and the radius \( R \) from the model generation volume = 4.57 E-5 cm, the number of generated carriers is calculated as:

\[
n(0) = 7.8E14 \text{ carriers/cm}^3
\]  

It is expected that the NOBIC carrier density is smaller than that of \( \tau \)-NSOM as calculated in Chapter 2, since the penetration depth of the NOBIC measurement is approximately one-tenth that of La Rosa’s \( \tau \)-NSOM experiment (300 nm vs \( 1/\alpha = 3.5 \mu m \)).

Figure 4.8 shows color versions of the NOBIC maps in Figure 4.7, and their corresponding background-subtracted topographical maps. The topographical maps are generated simultaneously with the NOBIC maps using the shear force feedback method described in Chapter 3. The line scans “A” and “B” from Figure 4.7 are again chosen to correlate the surface features to areas of NOBIC contrast. The conversion of the raw data to the photocurrent values is the same as given in Equation (4.7). The height values for the topographical map are determined using the formula:

\[
\text{Height, } Z = \frac{\text{raw value}}{32768} (10 \text{ V}) (\times 15 \text{ gain, control unit}) (\Delta z = 5.5 \text{ nm/V})
\]  

In addition to showing the effect of background signal fluctuations in the NOBIC image, line scan “B” is chosen for analysis because it highlights enhanced areas of NOBIC contrast compared to line scan “A”, as marked by the black regions of Figure 4.8a).
Comparing the topographical maps in Figure 4.8b) to the NOBIC images of Figure 4.8a), the NOBIC contrast is due to a line of defects, approximately 20 - 40 nm in height.

![Image](image_url)

Figure 4.8. Correlation of NOBIC contrast to surface topography. a) Color version of NOBIC images of Figure 4.7. b) Topography maps of NOBIC scan region. The height variations of the white lines, corresponding to “A” and “B”, are shown in Figure 4.8.

The surface characteristics are further detailed in Figure 4.9. The height measurement for line scans “A” and “B” is shown Figure 4.9a). As in the Ge mesa calibration measurement, the background of the topography image has “negative height”, due to the
background subtraction of the original image. Although negative height does not make physical sense, and can be alleviated by offsetting the height data by some amount so that the data reflects only positive values, the values are left as originally calculated after subtraction to preserve as much of the actual measurement data. The defect height is determined by subtracting the highest and lowest height values of the line scans, which are the values listed in the legend in Figure 4.9a).

A three dimensional projection of the topographical image is shown in Figure 4.9b). A ridge of defects appears in the image in yellow, with three prominent cones in the center of the ridge. Line scan “A” crosses the second defect from the left edge of the image, while line scan “B” crosses the rightmost of the three large defects. It is surmised that the NOBIC contrast seen in Figures 4.7 and 4.8 is due to carrier recombination with these defects.

Figure 4.9. Height measurements of line scans “A” and “B”. a) Two-dimensional line scans of “A” (pink), and “B” (purple) showing defect heights of 22 and 40 nm, respectively. b) Three-dimensional projection of Figure 4.8b), with defects associated with line scans “A” and “B” labeled.
A final word of caution must be given with respect to this analysis. As best as possible using the available techniques, a case has been made for the presence of intragrain defects in BP Solar polycrystalline silicon wafers. Furthermore, an argument has been given that shows that NOBIC contrast is caused by these defects on the surface of the material. In Figure 4.8, a line of discrete features, as seen in the topographical images, generates an unbroken line of NOBIC contrast, as opposed to discrete regions of contrast. However, in Figure 4.10a), a different line defect generates two unbroken lines of reduced photocurrent. At this stage, it is unclear whether the images presented in both these figures are due to intragrain defects, precipitate formation at these defects, or even surface contamination.

Figure 4.10 Maps of second line defect. a) NOBIC image, showing two lines of reduced photocurrent contrast. b) Surface topography, showing single line defect, surmised to be undecorated dislocation array.

It is unlikely that the surface features in these figures are due to surface contamination, such as dust, since such non-conducting material would generate little
photocurrent, and thus signal contrast, in the NOBIC images. The continuity of the line defect in Figure 4.10b) suggests an undecorated dislocation array. By contrast, the discrete nature of the line defect topography in Figure 4.8b) suggests precipitate formation along a dislocation array. The spacing between the lines of the NOBIC image in Figure 4.10a) is approximately 400 nm, while the full width at half maximum of the line defect NOBIC image in Figure 4.8 ranges from 400-600 nm. This suggests that the two lines of NOBIC contrast that surround the undecorated line defect blend together into one line due to recombination from the precipitate aggregation. Ultimately, the question still remains as to what surface and bulk defect interactions in the material generate the NOBIC contrast. Further analysis using transmission electron microscopy (TEM), X-ray diffraction (XRD), and deep level transient spectroscopy (DLTS) is required to fully ascertain the dislocation and impurity behavior within the BP Solar material.
4.4 REFERENCES


CHAPTER 5: CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

In this work, a near-field scanning optical microscope system, capable of imaging the recombination of defects in solar grade silicon, is established. A theoretical model, based on the work of Xu, et al, of the diffusion behavior of the photoexcited carriers using the convolution of carrier density and Gaussian light intensity to correlate tip illumination to the number of expected carriers available for NOBIC photocurrent or τ-NSOM free carrier absorption. Although the theoretical bases for the two imaging modes, based on photoinduced junction current (NOBIC) and infrared free carrier absorption (τ-NSOM), are described, images were generated only using the NOBIC mode.

Two types of NOBIC images were generated in this work. The first was an image whose contrast was caused by a line of defects approximately 20-40 nm in height. The second image was a two line image, whose contrast is due to a line of defects approximately 5-10 in height. The NOBIC contrast of the first image is believed to be due to precipitate formation along a dislocation array, while the contrast of the second image is due to an undecorated dislocation array.

5.2 Improvements and Additions

The necessary additions to this work are the τ-NSOM lifetime mapping results. The main design challenge to generate these maps is to find a way to pulse the visible beam at a
frequency comparable to the lifetime of the solar cell material. Based on the μPCD results (Chapter 4, Figure 4.1) of the BP Solar wafers, the highest lifetime values of a given wafer are about 20 μs. Since the infrared transmission is inversely proportional to the chopping frequency of the visible beam, a frequency of \( \frac{1}{20 \, \mu s} = 50 \, \text{kHz} \) is minimally required to image recombination effects in high lifetime regions, and will increase as lifetime decreases. The best mechanical choppers can reach 50 kHz, but they introduce too much oscillatory noise to the laser coupling to be effective. Instead, following the design of La Rosa, an acousto-optic modulator will be installed into the NSOM system in order to reach the required modulation frequencies.

Once the NOBIC and NSOM modes are fully established, the next step is to measure passivated and unpassivated samples to determine the role of passivation in improving the intragrain defect lifetime. Another future mode to establish is a reflectance mode, which measures the reflected light intensity from the sample surface. In such a way, reflectance can be used as a complement to the shear force topography measurement.
APPENDIX A: 1D CARRIER DIFFUSION MODEL
Given:

- \( n(x) \) = excess carrier density at lateral position \( x \)
- \( \tau(x) \) = carrier lifetime
- \( S(x) \) = carrier generation source with generation rate \( S_o2R \)
- \( a \) = defect size
- \( x_o \) = tip position

For carrier lifetime \( \tau(x) \) of:

\[
\tau(x) = \begin{cases} 
0, & 0 < x < a \\
\tau_o, & a < x < \infty 
\end{cases}
\]
Carrier diffusion equation is:

\[ n(x) = 0, \quad 0 < x < a \]

\[ \frac{d^2 n(x)}{dx^2} - \frac{n(x)}{D \tau(x)} + \frac{S(x)}{D} = 0, \quad a < x < \infty \]

rearranging…

\[ \frac{d^2 n(x)}{dx^2} - \frac{n(x)}{D \tau(x)} = -\frac{S(x)}{D} \]

where \( S(x) = S_0 \delta(x - x_o) \). Solving for homogeneous equation:

\[ n''(x) - \frac{n(x)}{D \tau} = 0 \]

homogeneous solution \( n_H \) is:

\[ n_H(x_I) = A \exp\left(x - x_o \sqrt{D \tau}\right) + B \exp\left(-x + x_o \sqrt{D \tau}\right), \quad 0 < x < a \]

\[ n_H(x_{II}) = C \exp\left(x - x_o \sqrt{D \tau}\right) + E \exp\left(-x + x_o \sqrt{D \tau}\right), \quad a < x < x_o \]

\[ n_H(x_{III}) = C \exp\left(x - x_o \sqrt{D \tau}\right) + E \exp\left(-x + x_o \sqrt{D \tau}\right), \quad x_o < x < \infty \]

and derivative…

\[ n'_H(x_I) = \frac{A}{\sqrt{D \tau}} \exp\left(\frac{x - x_o}{\sqrt{D \tau}}\right) - \frac{B}{\sqrt{D \tau}} \exp\left(\frac{-x + x_o}{\sqrt{D \tau}}\right), \quad 0 < x < a \]

\[ n'_H(x_{II}) = \frac{C}{\sqrt{D \tau}} \exp\left(\frac{x - x_o}{\sqrt{D \tau}}\right) - \frac{E}{\sqrt{D \tau}} \exp\left(\frac{-x + x_o}{\sqrt{D \tau}}\right), \quad a < x < x_o \]

\[ n'_H(x_{III}) = \frac{C}{\sqrt{D \tau}} \exp\left(\frac{x - x_o}{\sqrt{D \tau}}\right) - \frac{E}{\sqrt{D \tau}} \exp\left(\frac{-x + x_o}{\sqrt{D \tau}}\right), \quad x_o < x < \infty \]

Nonhomogeneous solution:

\[ n''(x) - \frac{n(x)}{D \tau} = -\frac{S(x) \delta(x - x_o)}{D} \]
If function \( f(x) \) is:

\[
f(x) = \begin{cases} 
0, & 0 < x < a \\
\exp(x - x_o / \sqrt{D\tau}), & a < x < x_o \\
\exp(-x + x_o / \sqrt{D\tau}), & x_o < x < \infty 
\end{cases}
\]

Particular solution \( n_p(x) \) is:

\[
n_p(x) = \int_{x_o}^{x} f(x) \, dx,
\]

\[
n_p(x_i) = \int_{x_o}^{x} 0 \, dx = 0, \quad 0 < x < a
\]

\[
n_p(x_H) = \int_{a}^{x} \exp(x - x_o / \sqrt{D\tau}) \, dx, \quad a < x < x_o
\]

\[
= \left[ \int_{a}^{x} \exp(x - x_o / \sqrt{D\tau}) \, dx \right]_{a=x_o}^{x}
\]

\[
n_p(x_{III}) = \int_{x_o}^{x} \exp(-x + x_o / \sqrt{D\tau}) \, dx, \quad x_o < x < \infty
\]

\[
= \left[ \int_{x_o}^{x} \exp(-x + x_o / \sqrt{D\tau}) \, dx \right]_{x=x_o}^{x}
\]

introducing step function into the derivative:

\[
n_p'(x_H) = \int \left( H(x-a) - H(x-x_o) \right) \left[ \exp(x - x_o / \sqrt{D\tau}) \right] \, dx, \quad a < x < x_o
\]

\[
n_p'(x_{III}) = \int H(x-x_o) \left[ \exp(-x + x_o / \sqrt{D\tau}) \right] \, dx, \quad x_o < x < \infty
\]

and:

\[
n_p''(x_H) = \int \left( \delta(x-a) - \delta(x-x_o) \right) \left[ \exp(x - x_o / \sqrt{D\tau}) / \sqrt{D\tau} \right] + \frac{I}{\sqrt{D\tau}} (H(x-a)H(x-x_o)) \exp(x - x_o / \sqrt{D\tau}), \quad a < x < x_o
\]

\[
n_p''(x_{III}) = \int \delta(x-x_o) \left[ \exp(-x + x_o / \sqrt{D\tau}) \right] + \frac{I}{\sqrt{D\tau}} H(x-x_o) \exp(-x + x_o / \sqrt{D\tau}) \, dx, \quad x_o < x < \infty
\]

combining terms for non-homogeneous solution:
total derivative $n'(x)$:

$$
n'(x) = n'_H(x) + n'_p(x);
$$

$$
n'(x_H) = \frac{A}{\sqrt{D\tau}} \exp \left( \frac{x - x_o}{\sqrt{D\tau}} \right) - \frac{B}{\sqrt{D\tau}} \exp \left( \frac{-x + x_o}{\sqrt{D\tau}} \right) = 0, \quad 0 < x < a
$$

$$
n'(x_H) = \frac{C}{\sqrt{D\tau}} \exp \left( \frac{x - x_o}{\sqrt{D\tau}} \right) - \frac{E}{\sqrt{D\tau}} \exp \left( \frac{-x + x_o}{\sqrt{D\tau}} \right) + \frac{L}{\sqrt{D\tau}} \exp \left( \frac{x - x_o}{\sqrt{D\tau}} \right), \quad a < x < x_o
$$

$$
n'(x_{III}) = \frac{C}{\sqrt{D\tau}} \exp \left( \frac{x - x_o}{\sqrt{D\tau}} \right) - \frac{E}{\sqrt{D\tau}} \exp \left( \frac{-x + x_o}{\sqrt{D\tau}} \right) - \frac{L}{\sqrt{D\tau}} \exp \left( \frac{-x + x_o}{\sqrt{D\tau}} \right), \quad x_o < x < \infty
$$
and second derivative \( n''(x) \):

\[
n''(x) = n_{II}''(x) + n_{III}''(x);
\]

\[
n''(x) = 0, \quad 0 < x < a
\]

\[
n''(x_{II}) = \frac{C}{D\tau} \exp\left(\frac{x-x_o}{\sqrt{D\tau}}\right) + \frac{E}{D\tau} \exp\left(\frac{-x+x_o}{\sqrt{D\tau}}\right) + \frac{I}{D\tau} \exp\left(\frac{x-x_o}{\sqrt{D\tau}}\right), \quad a < x < x_o
\]

\[
n''(x_{III}) = \frac{C}{D\tau} \exp\left(\frac{x-x_o}{\sqrt{D\tau}}\right) + \frac{E}{D\tau} \exp\left(\frac{-x+x_o}{\sqrt{D\tau}}\right) + \frac{I}{D\tau} \exp\left(\frac{-x+x_o}{\sqrt{D\tau}}\right), \quad x_o < x < \infty
\]

that is:

\[
n''(x) = 0, \quad 0 < x < a
\]

\[
n''(x_{II}) = \frac{1}{D\tau} n(x_{II}), \quad a < x < x_o
\]

\[
n''(x_{III}) = \frac{1}{D\tau} n(x_{III}), \quad x_o < x < \infty
\]

which satisfies the differential equation.

Boundary conditions:

\( n(x) \), \( n'(x) \) must be continuous at \( x = a \), and \( x = x_o \); \( n(x) = 0, \quad 0 < x < a \), \( n(\infty) = 0 \)

\[
n(x) = n(x_{II}), x = a:
\]

\[
n(x) = 0 = A \exp(x-x_o/\sqrt{D\tau}) + B \exp(-x+x_o/\sqrt{D\tau}); \quad \rightarrow A = 0, B = 0
\]

\[
0 = C \exp(a-x_o/\sqrt{D\tau}) + E \exp(-a+x_o/\sqrt{D\tau}) + I \exp(a-x_o/\sqrt{D\tau});
\]

\[
n(\infty) = 0:
\]

\[
C = 0
\]

\[
E \exp(-a+x_o/\sqrt{D\tau}) = -I \exp(a-x_o/\sqrt{D\tau})
\]

\[
E = -I \exp(2(a-x_o)/\sqrt{Dt})
\]
derivative continuity...
\[ f(x) = \frac{\exp(x - x_o)}{\sqrt{D \tau}} \quad a < x < x_o \]
\[ = \frac{\exp(-x + x_o)}{\sqrt{D \tau}} \quad x_o < x < \infty \]
\[ n(x) = I \cdot f(x) = -\frac{S(x)}{D}; \]
\[ n'(x_H) = I \cdot f'(x) = \frac{1}{\sqrt{D \tau}} \left[ \exp(x - x_o / \sqrt{D \tau}) \right] = \frac{I}{\sqrt{D \tau}} f(x); \quad a < x < x_o \]
\[ n'(x_{III}) = I \cdot f'(x) = -\frac{1}{\sqrt{D \tau}} \left[ \exp(-x + x_o / \sqrt{D \tau}) \right] = -\frac{I}{\sqrt{D \tau}} f(x); \quad x_o < x < \infty \]

so that \(n(x)\) is...
\[ n(x) = n_H + n_p \]
\[ n(x_I) = 0, \quad 0 < x < a \]
\[ n(x_H) = -I \exp(2a - x - x_o / \sqrt{D \tau}) + I \exp(x - x_o / \sqrt{D \tau}); \quad a < x < x_o \]
\[ n(x_{III}) = -I \exp(2a - x - x_o / \sqrt{D \tau}) + I \exp(-x + x_o / \sqrt{D \tau}); \quad x_o < x < \infty \]

where
\[ I = \frac{S(x) \sqrt{D \tau}}{2D} \]

To determine carrier density \(N(x)\), one must integrate each \(n(x)\) equation over each interval and add the results, that is:
\[ N(x) = \int_0^a n(x_I)dx + \int_a^x n(x_H)dx + \int_{x_o}^x n(x_{III})dx \]
Since \( n(x_1) = 0 \), the first integral is 0. For \( n(x_{II}) \) …

\[
\int_a^x n(x_{II}) \, dx = \int_a^x I \exp \left( \frac{x - x_o}{\sqrt{D\tau}} \right) \, dx - \int_a^x I \exp \left( \frac{2a - x - x_o}{\sqrt{D\tau}} \right) \, dx
\]

evaluating…

\[
\int_a^x n(x_{II}) \, dx = I \sqrt{D\tau} \left( 1 + \exp \left( \frac{2(a - x_o)}{\sqrt{D\tau}} \right) - 2 \exp \left( \frac{a - x_o}{\sqrt{D\tau}} \right) \right)
\]

for \( n(x_{III}) \) …

\[
\int_{x_o}^x n(x_{III}) \, dx = \int_{x_o}^x I \exp \left( \frac{-x + x_o}{\sqrt{D\tau}} \right) \, dx - \int_{x_o}^x I \exp \left( \frac{2a - x - x_o}{\sqrt{D\tau}} \right) \, dx
\]

evaluating…

\[
\int_a^x n(x_{III}) \, dx = I \sqrt{D\tau} \left( 1 - \exp \left( \frac{2(a - x_o)}{\sqrt{D\tau}} \right) \right)
\]

adding terms from the integrals gives:

\[
N(x) = 2I \sqrt{D\tau} \left( 1 - \exp \left( \frac{(a - x_o)}{\sqrt{D\tau}} \right) \right) = S(x) \tau \left( 1 - \exp \left( \frac{(a - x_o)}{\sqrt{D\tau}} \right) \right) \left[ H(x - a) + H(a - x) \right]
\]

where \( H(x) \) represents the Heaviside step function.
For τ-NSOM, the excess carrier density is measured at the tip position. In this case $x = x_o$ such that $n(x)$ is:

$$n(x) = n_H + n_P$$

$$n(x) = 0, \quad 0 < x < a$$

$$n(x_H) = \frac{S(x) \sqrt{D \tau}}{2D} \left(1 - \exp \left(\frac{2(a - x)}{\sqrt{D \tau}}\right)\right); \quad a < x < x_o$$

$$n(x_H) = \frac{S(x) \sqrt{D \tau}}{2D} \left(1 - \exp \left(\frac{2(a - x)}{\sqrt{D \tau}}\right)\right); \quad x_o < x < \infty$$

Integrating $n(x)$ in each region and adding the results gives:

$$N(x) = S(x) \tau \left[1 - \exp \left(\frac{2(a - x)}{\sqrt{D \tau}}\right)\right] \left[H(x - a) + H(a - x)\right]$$

Finally, we convolute the carrier density with a Gaussian distribution that represents the shape of the intensity distribution due to Fraunhofer diffraction from a circular aperture. This gives the carrier density for near-field light emanating from a subwavelength aperture.

For a normalized Gaussian distribution $G(x_o - x)$:

$$G(x_o - x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(x_o - x)^2}{2\sigma^2}\right)$$

and $N(x)$ as given above, the convolution $C(x)$ is given as:

$$C(x) = \int_{-\infty}^{\infty} G(x_o - x)N(x)dx$$

For NOBIC, $C(x)$ is:

$$C(x) = \frac{S(x) \tau}{\sqrt{2\pi\sigma^2}} \left[\int_a^\infty \exp \left(-\frac{(x_o - x)^2}{2\sigma^2}\right) \left(1 - \exp \left(\frac{a - x}{L}\right)\right) dx + \int_{-\infty}^a \exp \left(-\frac{(x_o - x)^2}{2\sigma^2}\right) \left(1 - \exp \left(\frac{a - x}{L}\right)\right)\right]$$

where $L = \sqrt{D \tau}$. Changing variables in the second term gives…
\[ C(x) = \frac{S(x) \tau}{\sqrt{2 \pi \sigma^2}} \left[ \int_a^x \exp \left( \frac{-(x_a - x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{a - x}{L} \right) \right) \, dx + \int_x^\infty \exp \left( \frac{-(x_a + x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{a - x}{L} \right) \right) \, dx \right] \]

so that …

\[ C(x) = \frac{S(x) \tau}{\sqrt{2 \pi \sigma^2}} \left[ \int_a^x \exp \left( \frac{-(x_a - x)^2}{2\sigma^2} \right) + \exp \left( \frac{-(x_a + x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{a - x}{L} \right) \right) \, dx \right] \]

this is a four term integral which ultimately yields…

\[ C(x) = \frac{S(x) \tau}{2\sqrt{2 \pi \sigma^2}} \left[ \text{erfc} \left( \frac{a - x_a}{\sqrt{2\sigma}} \right) + \text{erfc} \left( \frac{a + x_a}{\sqrt{2\sigma}} \right) - \exp(\Phi_1) \text{erfc}(\phi_1) - \exp(\Phi_2) \text{erfc}(\phi_2) \right] \]

where…

\[ \Phi_1 = \left( \frac{a - x_a}{L} + \frac{\sigma^2}{2L^2} \right), \quad \phi_1 = \left( \frac{a - \xi}{\sqrt{2\sigma}} \right), \quad \xi = x_a - \left( \frac{\sigma^2}{L} \right) \]
\[ \Phi_2 = \left( \frac{a + x_a}{L} + \frac{\sigma^2}{2L^2} \right), \quad \phi_2 = \left( \frac{a + \mu}{\sqrt{2\sigma}} \right), \quad \mu = x_a + \left( \frac{\sigma^2}{L} \right) \]

For \( r \)-NSOM, \( C(x) \) is:

\[ C(x) = \frac{S(x) \tau}{\sqrt{2 \pi \sigma^2}} \left[ \int_a^x \exp \left( \frac{-(x_a - x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{2(a - x)}{L} \right) \right) \, dx + \int_x^\infty \exp \left( \frac{-(x_a - x)^2}{2\sigma^2} \right) \left( 1 - \exp \left( \frac{2(a - x)}{L} \right) \right) \, dx \right] \]

Following the above procedure, \( C(x) \) is given as:

\[ C(x) = \frac{S(x) \tau}{2\sqrt{2 \pi \sigma^2}} \left[ \text{erfc} \left( \frac{a - x_a}{\sqrt{2\sigma}} \right) + \text{erfc} \left( \frac{a + x_a}{\sqrt{2\sigma}} \right) - \exp(\Phi_1) \text{erfc}(\phi_1) - \exp(\Phi_2) \text{erfc}(\phi_2) \right] \]

where…

\[ \Phi_1 = \left( \frac{2(a - x_a)}{L} + \frac{\sigma^2}{2L^2} \right), \quad \phi_1 = \left( \frac{a - \xi'}{\sqrt{2\sigma}} \right), \quad \xi' = x_a - \left( \frac{2\sigma^2}{L} \right) \]
\[ \Phi_2 = \left( \frac{2(a + x_a)}{L} + \frac{\sigma^2}{2L^2} \right), \quad \phi_2 = \left( \frac{a + \mu'}{\sqrt{2\sigma}} \right), \quad \mu' = x_a + \left( \frac{2\sigma^2}{L} \right) \]
APPENDIX B: SHEAR FORCE FEEDBACK GAIN
Given feedback loop consisting of:

\[ V_{\text{ref}} = \text{Reference voltage from current setpoint dial, nanoscope control unit} \]
\[ V_{\text{out}} = \text{Output voltage after amplification} \]

where:

- \( V_{\text{ref}} \):
  - Reference voltage from current setpoint dial, nanoscope control unit

- \( V_{\text{out}} \):
  - Output voltage after amplification

- \( G_{\text{PI}} \):
  - Proportional + Integral gain from control unit

- \( G_{\text{HV}} \):
  - High voltage gain, control unit

- \( G_{\text{PZT}} \):
  - Piezo electric tube gain = 5.5 nm/volt, Chapter 3

- \( G_{\text{PRE}} \):
  - Tuning fork preamplifier gain

- \( G_{\text{LIA}} \):
  - Lock in amplifier gain = 1 / 4\( \tau \), \( \tau \) = time constant

Closed loop gain:

\[
\frac{V_{\text{out}}}{V_{\text{ref}}} = \frac{G}{1 + G}
\]

Open loop gain, \( G \):

\[
G = G_{\text{PI}} \times G_{\text{HV}} \times G_{\text{PZT}} \times G_{\text{PRE}} \times G_{\text{LIA}}
\]

**Proportional + Integral gain, \( G_{\text{PI}} \):**

Circuit diagram:
Gain $R_5 / R_1$: from block diagram:

\[
\frac{R_5}{R_1} = \frac{-Z_2}{Z_1} = -\frac{100\,\text{k}\Omega}{100\,\text{k}\Omega} = -1
\]

\[
\frac{R_3}{R_2} = \frac{Z_5}{Z_2 + Z_5} = \frac{\frac{1}{j\omega C}}{R + \left(\frac{1}{j\omega C}\right)} = \frac{1}{R + (j\omega C + 1)} = \frac{1}{(j\omega \omega_o) + 1} = \frac{1}{\sqrt{1 + (\omega/\omega_o)^2}}, \quad \omega_o = \frac{1}{RC}
\]

\[
\frac{R_4}{R_2} = \frac{-Z_6}{Z_4} = -\frac{100\,\text{k}\Omega}{101\,\text{k}\Omega} = -.990
\]

\[
R_5 = \frac{-Z_2}{Z_1} R_3 - \frac{-Z_2}{Z_1} R_4
\]

\[
\frac{R_5}{R_1} = \frac{-Z_2}{Z_1} \frac{R_3}{R_1} - \frac{-Z_2}{Z_1} \frac{R_4}{R_1} = \frac{-2k\Omega}{2k\Omega} \frac{R_3}{R_1} - \frac{-2k\Omega}{100k\Omega} \frac{R_4}{R_1} = \frac{-R_3}{R_1} - \frac{.02R_4}{R_1} = \frac{-R_3}{R_1} - \frac{.0198R_4}{R_1} = -\frac{R_3}{R_1} + .0198
\]

\[
= \frac{1}{\sqrt{1 + (\omega/\omega_o)^2}} + .0198
\]
**High voltage gain, $G_{HV}$:**

Circuit Diagram:

![Circuit Diagram](image)

Block diagram:

![Block Diagram](image)

Gain, $R_7 / R_5$:

$$R_6 = \frac{Z_{12} Z_{11}}{Z_{10}} R_5,$$

$$R_7 = Z_{42} R_6 = - \frac{Z_{12} Z_{11}}{Z_{10}} R_5,$$

$$\frac{R_7}{R_5} = - \frac{Z_{12} Z_{11}}{Z_{10}}$$

$$= \left( \frac{10 k\Omega + 10 k\Omega}{100 k\Omega} \right) \frac{30 k\Omega}{2 k\Omega} = -15$$
Tuning fork preamplifier gain, $G_{\text{PRE}}$:

Circuit Diagram:

![Circuit Diagram](image)

Block Diagram:

![Block Diagram](image)

$$\frac{R_9}{R_8} = -\frac{Z_{14}}{Z_{13}} = -\frac{R_f}{R_{in}+(1/j\omega C)} = \frac{(\omega R_f C)^2}{1-(\omega R_{in} C)^2} = \frac{(\omega_1/\omega_f)^2}{1-(\omega_1/\omega_{in})^2},$$

$$\omega_f = \frac{1}{R_f C} = \frac{1}{(10\text{M}\Omega)(0.03\text{nF})} = 3.33 \text{ MHz}, \quad \omega_{in} = \frac{1}{R_{in} C} = \frac{1}{(1\text{k}\Omega)(0.03\text{nF})} = 33.3 \text{ GHz}$$

$$\frac{R_{\text{out}}}{R_9} = \frac{Z_{16} + Z_{17}}{Z_{15}} = -\frac{100.01k\Omega}{10k\Omega \cdot 4.7k\Omega} = -31.35$$
Gain-Frequency Plot:

Gain vs. Frequency graph with the following labels:

- Gain
- Frequency

- $G_{\text{INT}}$
- $G_{\text{PROP}}$
- $6 \text{ dB/oct.}\ (1/f)$
- $60 \text{ dB/oct.}\ (\text{LIA})$

Equation for $6 \text{ dB/oct.}\ (1/f)$:

$$\frac{1}{4\tau}$$