Fabrication and Characterization of Nanoantennas

by
Christopher Rohde

A thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg

UNIVERSITY OF MANITOBA

Copyright © 2006 by Christopher Rohde

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Fabrication and Characterization of Nanoantennas

BY

Christopher Rohde

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

OF

MASTER OF SCIENCE

Christopher Rohde © 2006

Permission has been granted to the Library of the University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.

This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.

Abstract

The utility of light in science and engineering is unquestionable, yet it faces a severe limitation as the size of many objects of interest has shrunk below the wavelength of light. The Abbé diffraction limit[1] tells us that the use of light is impeded in investigating and interacting with subwavelength features. To overcome this limit, we look to near-field techniques. Abandoning the far-field approach and allowing close interaction between the light source and the sample can achieve sub-wavelength resolution. There are several near-field techniques. The method examined here is based on locally enhancing electric fields by introducing metallic objects into optical beams in order to locally enhance the field strength. We have used an optical fiber to couple optical beams into proximity of an antenna structure to convert the optical beam to a localized electric field. The antenna structure used is a bowtie antenna, which was selected due to its relative ease of fabrication. Bowtie antennas have been demonstrated to be effective near-field couplers at 458 nm[2] and 830 nm[3][4] wavelengths.

Our investigation is focused on the near-infrared regime, specifically 1550 nm. We have simulated bowtie antennas using 3-D electromagnetic modeling tools. Bowtie antennas of Ag, Au, Cu and Al were simulated for film thicknesses of 40 nm. With a 10 nm gap simulations predicted electric fields enhanced by factors of as much as 140$\times$ at wavelengths of 1550 nm. Test structures have been fabricate using two different methods: optical lithography combined with focused ion beam etching, and electron beam lithography. Local electric fields were imaged using heterodyned electric force microscopy as well as scattered near-field detection. Results from the latter technique indicate near-field coupling is occurring and the field pattern has some of the features observed in the simulations. In addition, field enhancements were observed from features produced by the lift off fabrication that are not associated with the gap. This observation suggests caution in interpreting field enhancement measurements without field mapping. The very high amplification predicted by the simulations in the gap region has yet to be successfully imaged.

Contents

1 Background/Motivation **1**
1.1 The Diffraction Limit 1
1.2 Near-field Scanning Optical Microscopy (NSOM) 3
1.2.1 Apertured NSOM 4
1.2.2 Apertureless NSOM 9
1.2.3 Scanning Tunneling Optical Microscopy (STOM) 11
1.3 Nanoantennas 12
1.4 Outline 19

2 Simulation **24**
2.1 Introduction 24
2.2 Material Properties 24
2.3 Simulation Geometry 29
2.4 Results 31
2.4.1 Basic Simulation 31
2.4.2 Parameter Variation 37
2.5 Conclusions 38

3 Antenna Layout and Illumination Apparatus **43**
3.1 Introduction 43
3.2 Illumination and Device Layout 43
3.3 Illumination Apparatus 45
3.4 Laser alignment 48
3.5 Scattered Field Detection Apparatus 50

4 Fabrication **54**
4.1 Introduction 54

4.2 Materials . . . 54
4.3 Deposition . . . 56
4.4 Device Layout Consideration . . . 56
4.5 Focused Ion Beam (FIB) Milling . . . 57
4.5.1 FIB Technology . . . 57
4.5.2 FIB Fabrication procedure . . . 61
4.5.3 FIB Fabrication results . . . 66
4.6 Electron Beam Lithography (EBL) . . . 68
4.6.1 EBL Technology . . . 68
4.6.2 EBL Fabrication procedure . . . 69
4.6.3 EBL Fabrication results . . . 75

5 Experiment **77**
5.1 Introduction . . . 77
5.2 Near-Field Detection via Force Measurement . . . 77
5.2.1 Apparatus . . . 77
5.2.2 Results . . . 81
5.3 Near-Field Detection via Scattered-Field Measurement . . . 85
5.3.1 Apparatus . . . 85
5.3.2 Results . . . 87

6 Conclusions **89**

A Fourier optics and optical limits **91**
A.1 Fourier Optics . . . 91
A.2 Light in Free space . . . 92
A.3 Approximation methods . . . 93
A.3.1 The Fresnel Approximation . . . 93
A.3.2 The Fraunhofer approximation . . . 95
A.4 Diffraction . . . 96
A.4.1 Fraunhofer Diffraction . . . 97
A.4.2 The Fraunhofer diffraction limit . . . 99

List of Tables

2.1 Dielectric Loss Tangent of several metals at 1550 nm and 830 nm. 25

3.1 Specifications for Thor Labs 50-1550-FC grin fiber collimator. . . 46
3.2 Specification for Thor Labs A390-C molded glass aspheric lens. 46
3.3 Typical detector measurements for 100 μW input laser power. . 50
3.4 Specification for Thor Labs A397TM-C molded glass aspheric lens. 50
3.5 Specifications for Thor Labs FGA21 InGaAs-Photodiode. 51

4.1 Raith 150 Specification. 69

5.1 CSCS12 contact mode probe specifications. 78
5.2 PPP-NCH tapping mode probe specifications. 86

List of Figures

1.1 Tip/Sample interaction in various NSOM schemes. 4
1.2 Formation of tapered fiber tips by laser heating and pulling. . . 6
1.3 Block diagram of a typical apertured NSOM apparatus. 7
1.4 Geometries of apertureless tips modeled in [24]. 10
1.5 Basic apparatus use in apertureless NSOM. 10
1.6 Basic apparatus for PSTM/STOM. 12
1.7 3.2 GHz scale model illumination, measurement and antenna apparatus. 14
1.8 Field intensity measured from scale model in [32]. 15
1.9 Maskless method for fabrication of bowtie antennas of a pyramidal cantilevered tip. 15
1.10 Near-field measurement of pyramidal antenna. 16
1.11 TPPL excitation of bowtie array. 17
1.12 Comparison of theoretical versus measured field enhancement versus gap width. 18
1.13 Nanoantennas and WLSC generation. 20
1.14 Measured WLC spectra for different excitation powers. 21
1.15 WLC power versus average incident laser power 21
1.16 Simulated and measured relationship of WLC power versus antenna length. 22
1.17 SEM image of aperture array. 22
1.18 NSOM image of aperture array. 23

2.1 Relationship of ϵ_r versus wavelength. 26
2.2 Relationship of ϵ_i versus wavelength. 27
2.3 Relationship of the dielectric loss tangent (ϵ_i/ϵ_r) versus wavelength. 28

2.4 Simulation convergence data. 31
2.5 Electric field inside the metal arms of the antenna. 32
2.6 Electric-field along the x-axis at the top of the antenna 1 nm from the metal, given a 1 V/m linearly polarize plane wave excitation. 33
2.7 Electric-field along the y-axis at the top of the antenna 1 nm from the metal, given a 1 V/m linearly polarize plane wave excitation. 34
2.8 Electric-field versus distance away from antenna, given a 1 V/m linearly polarize plane wave excitation. 35
2.9 Electric field in the plane above the antenna. 36
2.10 Magnetic field in the plane above the antenna. 36
2.11 Electric field in the plane above an antenna with connected halves. 36
2.12 Magnetic field in the plane above an antenna with connected halves. 37
2.13 Amplification factor versus antenna length. 38
2.14 Amplification factor versus gap size. 39
2.15 Amplification factor versus metal thickness. 40
2.16 Amplification factor versus flare angle. 41
2.17 Amplification factor versus excitation wavelength for different metals. 42

3.1 Components required for beam alignment. 44
3.2 Illustration of the tapering channel and its use for beam alignment. 45
3.3 Individual apparatus components. 47
3.4 Alignment apparatus. 47
3.5 Procedure for gluing samples to ensure laser alignment. 49
3.6 Spectral Response for FGA21 photodiode. 51
3.7 Focusing scattered field detector. 52
3.8 Illumination apparatus combined with scattered field detector. . 53

4.1 Desired antenna geometry. 55
4.2 Components of ion column. 58
4.3 Modes of operation of the focused ion beam. 59
4.4 Fabrication with focused ion beam. 61
4.5 Mask Layout. 63

4.6 Optical micrograph of right side window and tapering channel at 10× and 100× magnification. 63
4.7 Optical micrograph of left side window and tapering channel at 10× and 100× magnification. 64
4.8 Illustration of beam alignment in a FIB system. 65
4.9 Order of Ion beam milling. 65
4.10 Results of focused ion beam fabrication. 67
4.11 First posible EBL fabrication technique. 70
4.12 Exposure test of HSQ resist. 71
4.13 Second posible EBL fabrication technique. 72
4.14 Nanoantennas fabricated using electroplating. 72
4.15 Final EBL fabrication technique for creating the antennas. 73
4.16 Final EBL fabrication technique for creating the tapering channels and windows. 74
4.17 Results of electron beam fabrication. 76

5.1 Block Diagram for detection though force measurement. 79
5.2 Digital Instruments Enviroscope. 80
5.3 Nanoantenna force measurement results. 82
5.4 Simulation results of fields along the x-axis 50 nm above the antenna. 83
5.5 Simulation results of fields along the y-axis 50 nm above the antenna. 84
5.6 Block Diagram for detection though scattered field measurement. 86
5.7 Nanoantenna scattered-field measurement results. 87

A.1 Free space transfer function. 92
A.2 Fraunhofer diffraction from a circular aperture. 98

List of Copyrighted Material for which Permission was Obtained

1.7 3.2 GHz scale model illumination, measurement and antenna apparatus. Reused with permission from Robert D. Grober, Robert J. Schoelkopf, and Daniel E. Prober, Applied Physics Letters, 70, 1354 (1997). Copyright 1997, American Institute of Physics. . . . 14

1.8 Field intensity 2.5 cm from waveguide opening. Reused with permission from Robert D. Grober, Robert J. Schoelkopf, and Daniel E. Prober, Applied Physics Letters, 70, 1354(1997). Copyright 1997, American Institute of Physics. 15

1.11 TPPL excitation of bowtie array. Reprinted figure with permission from [3]. Copyright 2005 by the American Physical Society. 17

1.12 Comparison of theoretical versus measured field enhancement versus gap width. Reprinted figure with permission from [3]. Copyright 2005 by the American Physical Society. 18

1.13 Nanoantennas and WLSC generation. Reprinted with permission from AAAS. 20

1.14 Measured WLC spectra for different excitation powers. From [4]. Reprinted with permission from AAAS. 21

1.15 WLC power versus average incident laser power From [4]. Reprinted with permission from AAAS. 21

1.16 Simulated and measured relationship of WLC power versus antenna length. From [4]. Reprinted with permission from AAAS. 22

1.17 SEM image of aperture array. Reused with permission from [2]. Copyright 2006, American Institute of Physics. 22

1.18 NSOM image of aperture array. Reused with permission from [2]. Copyright 2006, American Institute of Physics. 23

Chapter 1

Background/Motivation

1.1 The Diffraction Limit

The utility of light in science and engineering is unquestionable, yet as technology progresses its use faces a severe limitation. In the past several decades the size of many objects of interest has shrunk to the point where they are smaller than the wavelength of light. This renders them inaccessible to traditional optical microscopy techniques. A traditional optical microscope illuminates the sample with an electromagnetic wave which interacts with the device under test. The scattered light (either reflected or transmitted) is collected through the use of a lens and then sent to a detector. The reason that subwavelength features are inaccessible to these methods is that the light scattering from the high-resolution features of the sample occurs in such a way that the field decays exponentially[1]. The maximum theoretical resolution of an optical microscope of this type is typically given as the Abbé diffraction limit [1]

$$\Delta x = 0.61\lambda/NA, \tag{1.1}$$

where NA is the numerical aperture of the lens.

As an example, for a typical NA of 0.25, for light with λ = 400 nm the mini-

[1] A more detailed explanation of this process, and the derivation of the Abbé diffraction limit, can be found in Appendix A

mum distance between features that can be resolved is

$$\Delta x = 0.61\lambda/NA = 0.61\frac{400\times10^{-9}}{0.25} \approx 976\,\text{nm}.$$

Thus even in the theoretical limit, the smallest features that are resolvable are still nearly a micron in size.

To overcome the diffraction limit, a variety of techniques have been explored. The simplest solution is to employ a higher-refractive index material in the region between the sample and the objective. In term of the index of refraction, the numerical aperture can be calculated as:

$$NA = n\sin\theta, \tag{1.2}$$

where n is the index of refraction of the medium in which the lens is operating, and θ is one half the angular aperture of the lens. The best immersion oil lenses have an $NA \approx 1.5$, so even in the best case the resolution improvement from this technique is small. If we were to replace our objective in the earlier example with an immersion oil lens having $NA = 1.5$, we have

$$\Delta x = 0.61\lambda/NA = 0.61\frac{400\times10^{-9}}{1.5} \approx 163\,\text{nm}.$$

While this is significantly smaller than the limit of a standard objective, it is still far greater than the resolution desired.

Another obvious technique for reducing the minimum resolvable feature size is to use a shorter wavelength of light. This method has also been employed to improve the resolution of optical lithography, but it again suffers from several limitations. Designing optics to work in the UV regime is very challenging. This is because many materials are opaque to UV light, and can even damage the materials they are examining. Since one of the greatest benefits to optical analysis is that it is non-destructive, this is a significant limitation. Another significant restriction is that the spectroscopic information available in the UV regime is limited when compared to that of the visible or near IR, and to get this information the wavelength would have to be increased. With short-wavelength analysis none of this information can be measured locally. Finally, the resolution improvement is again small, and the use of this tech-

nique is restricted to certain specialized applications.

A common tool to approach the diffraction limit is the confocal microscope. The device is similar to a traditional optical microscope, but the confocal microscope employs a very tightly-focued spot to illuminate the sample. This spot is scanned over the sample and the image created by combining the locally illuminated points. This technique maintains many of the advantages of traditional optical microscopy, but again does not seem likely to offer resolution much below a wavelength. One advantage of confocal microscopy is that it can be applied to samples that require optical sectioning[5].

While the aforementioned techniques offer some improvement over traditional optics, we have seen that they all fail to overcome the diffraction limit beyond an order of magnitude. For greater resolution improvements, we must turn to another class of techniques.

1.2 Near-field Scanning Optical Microscopy (NSOM)

Currently, the most effective methods for subwavelength imaging are the family of techniques known as Near-field Scanning Optical Microscopy (NSOM), sometimes referred to as Scanning Near-field Optical Microscopy (SNOM). NSOM techniques disregard the far-field assumption made in traditional optics and instead interact with the near-field of the sample. The idea behind near-field imaging was first proposed in 1928 by the Irish scientist E. Synge [6]. Although he later disregarded the scheme, his idea of illuminating a sample using a small aperture in close proximity was entirely valid. Had Synge tried to realize his idea, he would have met with significant difficulty with the implementation. It was not until the development of the scanning tunneling microscope in 1984 that the technology to position the near-field source became available. NSOM was simultaneously re-discovered and developed shortly following this by the group of D. Pohl at the IBM Rüschilikon Research Laboratory[7][8] as well as by A. Lewis and E. Betzig at Cornell[9][10][11][12].

NSOM techniques typically fall in to two categories: Apertured NSOM (ANSOM) and apertureless NSOM. An interesting technique that combines features of both methods is the technique known as Scanning Tunneling Optical Microscopy (STOM). The general tip/sample interaction of these three

methods is shown in figure 1.1.

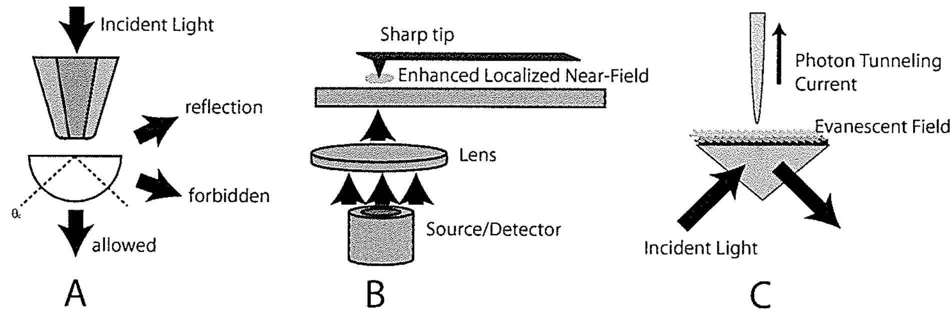


Figure 1.1: Tip/Sample interaction in various NSOM schemes. A) Apertured NSOM B) Apertureless NSOM C) Scanning Tunneling Optical Microscopy.

1.2.1 Apertured NSOM

The first NSOM technique developed is also currently the most common one. Apertured NSOM is closest to Synge's original method, and uses an illuminated subwavelength aperture to provide the strongly localized light source. Here strongly localized means that the illumination is mostly confined to the near-field. This source is brought in close to the sample and scanned across while the interaction is measured. This can be done with a far-field detector, or by using the probe in collection mode (the former is the more common technique). Perhaps the most challenging aspect of ANSOM is the fabrication of the apertured probe. There are two major difficulties in the construction of these probes: creating a sharp tip in an optical waveguiding material, and metallization and fabrication of the subwavelength aperture.

Apertured probe fabrication

The majority of techniques for fabricating NSOM probes fall into two categories: chemical etching or laser heating and pulling. There do exist other methods, but they are in the minority.

A review of chemical etching techniques can be found in [13]. The most successful chemical etching technique involves dipping a bared fiber in HF. In this process the fiber is dipped in a two-phase solution consisting of HF and an organic solvent. The difference in surface tension between the acid and the

solvent creates a meniscus, which causes the tip formation. The products of the etching flow downwards since they are heavier than the HF, and this flow will affect the etch rate at different depths. The freshly-etched regions will be filled with a combination of acid and etch products, and this increases the weight of the meniscus. As the reaction continues, the weight of the liquid will increase beyond that which can be supported by the surface tension, and the level of the liquid drops. This process continues, creating a tapering effect until the meniscus reaches the end of the fiber.

Fiber etching techniques fall into two categories: static etching and dynamic etching. Static etching methods keep the fiber stationary, and the meniscus is what dictates the etch geometry. The tip shape can be controlled by the choice of different solvents which will affect the surface tension. In dynamic etching meniscus techniques are still employed, but the fiber is moved vertically during the process. This can allow for greater control over the lateral etch rates at different vertical locations[13]. Additionally, while static etching is a self-terminating process, dynamic etching requires intervention to halt the process.

The other major category of tip manufacturing techniques is laser-heated pulling. A simple schematic of this process is shown in Figure 1.2. A fiber is clamped at both ends, and a laser beam is focused on to a stripped region. This is typically done using a CO_2 laser operating at 10.6 μm. This heats a small region, which softens that area. When a pulling force is applied a sharp tip is formed. The pulling force in a commercial puller is generated using solenoids. The laser power, spot size, dwell time, and the pulling force can all be varied to change the profile of the resulting tips.

A common requirement for NSOM tips is that the fiber be bent to suit some particular application. Again this can be accomplished via laser heating. One difference is that when bending a probe the laser is typically more tightly focused, as extra care must be taken to preserve the sharp point at the tip. One end of the fiber is fixed, and the laser is pointed at the desired bend point, typically 1 mm from the tip. The resulting bend angle is a product of the exposure time and the laser power.

Regardless of the process used in fabricating the pointed tip, the next step is to apply a metal coating. This is typically done via vacuum evaporation, with

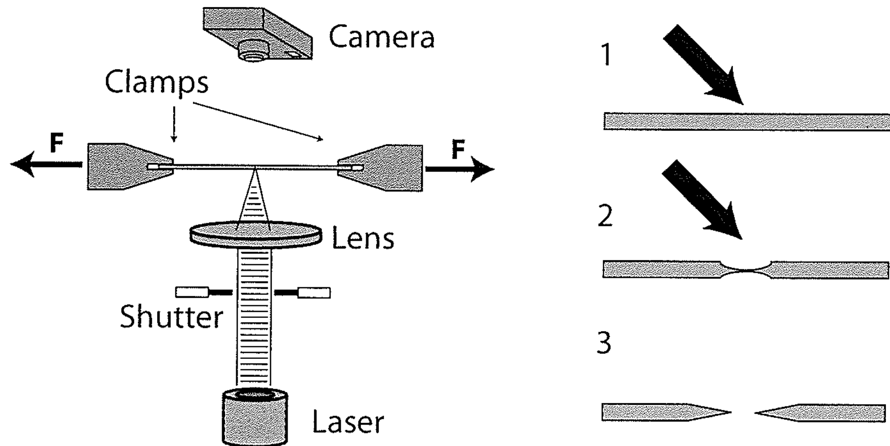


Figure 1.2: Formation of tapered fiber tips by laser heating and pulling.

the aperture being created using a shadow masking technique. The tips are mounted on a rotating chuck with the ends points upwards at a roughly 45° angle. The combination of rotation and mounting angle cause a small aperture to be created at the tip as the rest of the probe is coated with between 30-50 nm of metal.

Apertured NSOM apparatus

Figure 1.3 shows a block diagram of the components in a typical ANSOM system, this one using an external far field collection system. The illumination module begins with a laser with polarization control being focused in to the apertured probe. The tip is mounted in a support that enables coarse positioning and detection of the shear force. The scanning and collection module is used to control the distance between the tip and the sample, and to collect the illumination signal and channel is to the detector. The subwavelength resolution occurs because the aperture probe is designed so that it emits light mostly confined to the near field. This interacts with a subwavelength region of the sample, and though the interaction is measured in the far-field, the light that is detected is primarily a result of absorption, phase shifting, or locally excited fluorescence in the locally illuminated region[5].

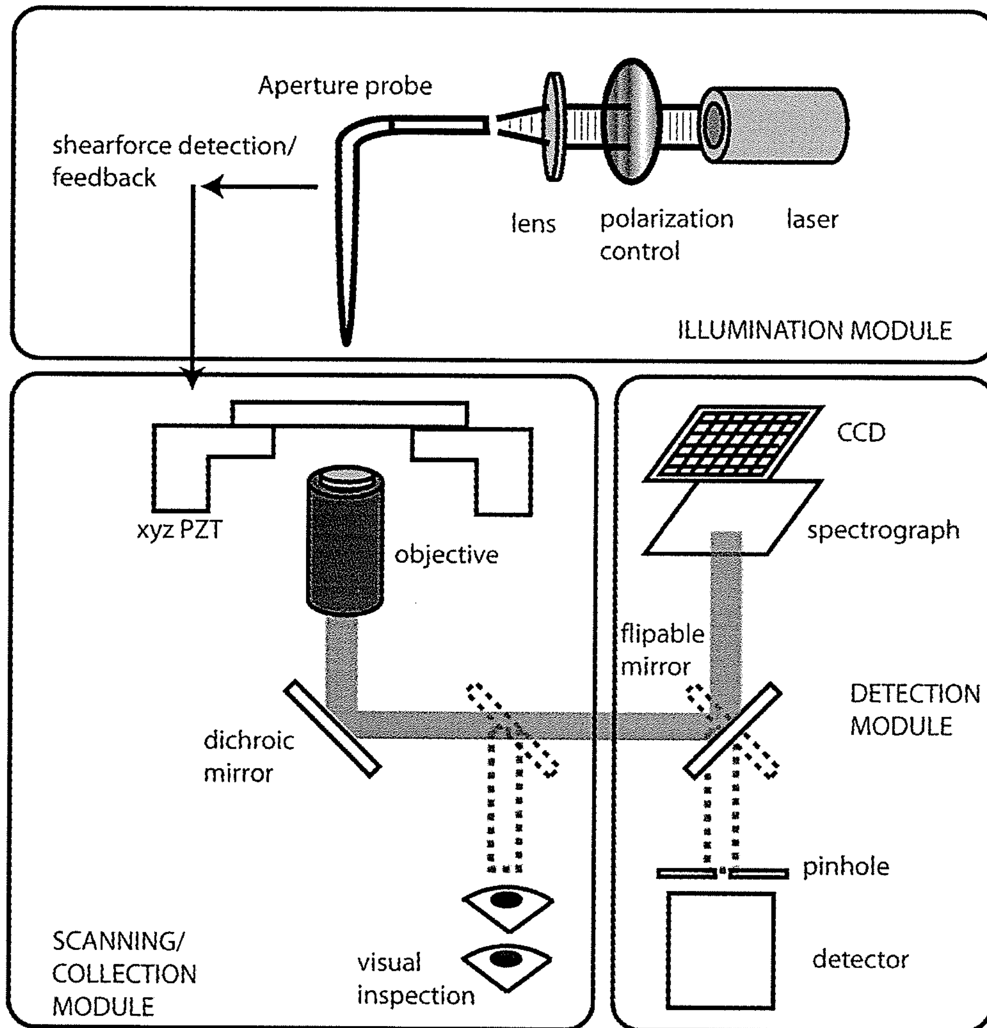


Figure 1.3: Block diagram of a typical apertured NSOM apparatus.

Advantages and Disadvantages of Apertured NSOM

The principle advantages of ANSOM lie in its improved resolution and the fact that most of the contrast mechanisms that occur have counterparts in traditional optical microscopy[5]. The importance of having similar contrast mechanisms cannot be overstated. This greatly simplifies interpretation of the results, as well as obviates the need for specially designed calibration samples. By virtue of this similarity, the extension of most optical techniques to these super resolutions becomes a trivial matter. One of the areas where this was beneficial is in the use of fluorescence staining techniques. In this technique ANSOM is useful because the strongly localized near-field source provides very little in the way of background radiation, which prevents photobleaching[5]. With ANSOM sub-100 nm resolution is easily achievable, and resolutions as low as 30 nm have been reported [14]. This resolution is 5-15x greater than a confocal microscope with 1.4 NA[5].

There are several limitations that restrict the usefulness of ANSOM. One of the primary issues is in the tip manufacturing process. The current state of the art gives relatively low yield, with a large variation between samples. As of this writing there are few commercial ANSOM tips available[2]. Even with manufacturing improvements, there are still certain fundamental limits inherent in the apertured probe concept. The most critical are the fact that the probe resolution has an ultimate limit, and that the near-field has a relatively low energy density. ANSOM has a resolution limit of between 30-50 nm[15], depending on the probe. When combined with fluorescence imaging, this can potentially be as low as 10 nm[16]. The ultimate resolution depends on the aperture size, but as this decreases, so does the amount of light that passes through the aperture, by a factor of a^{-4}[5], where a is the diameter of the aperture . Even for a probe with a 100 nm aperture, the throughput is typically only 10^{-4}[17].

Keeping in mind the fundamental resolution limit, there are still many objects of interest that can be investigated with NSOM that are not accessible with traditional optical techniques. In this regime, the second limitation of ANSOM prevents the use of many techniques. In order to perform localized spectroscopic analysis of samples, we would like a high energy density as well as have accessible a wide range of wavelengths. Similarly, to enable localized

[2]One vendor is Nanonics Imaging Ltd (www.nanonics.co.il)

material processing a high power density is desired. These limitations motivate the investigation of other near-field techniques.

1.2.2 Apertureless NSOM

The limitation of apertured NSOM come about from the key component of the system, the apertured tip. One technique that overcomes the probe limitations discussed above is apertureless NSOM[18]. This technique works by focusing a far-field source on to a sharply pointed probe tip. This results is a highly localized near-field concentration at the apex of the probe, and this signal is used as the illumination source. The interaction is once again collected in the far-field. This technique overcomes many of the disadvantages of ANSOM, but has its own drawbacks. These are discussed further in section 1.2.2.

Design and Fabrication of Apertureless NSOM probes

While the fabrication of an apertureless probe would seem to be a simple proposition, it is still advantageous to consider the optimal probe geometry for high near-field enhancement. If we model the tip as a sphere, then an analytical solution can be determined for the interaction between the tip and a plane wave[19]. Other geometries that can be solved analytically are pairs of spheres[20] and spheroids[21]. A single metal sphere on the end of a dielectric tip has been shown to be sufficient to generate near-field images below the diffraction limit[22][23]. These arrangements yield enhancements far below other arrangements, however.

To determine the optimal field enhancement, numerical methods are employed. Through Finite-Difference Time Domain (FDTD) simulations the tip enhancement of various configuration can be modelled. Figure 1.4 shows the tip geometries analyzed in [24]. The results of the simulation showed that for tips illuminated by 825 nm plane wave, the near-field at the tip of the quasi-infinite cone (Figure 1.4a) had an intensity enhancement factor of 361. The finite tip, however (Figure 1.4b), had an enhancement factor of 918. Even without any consideration of design, a much higher enhancement factor (roughly 2.5x) was obtained, due to resonance effects in the gold tip. By tuning the particle length and incident wavelength, enhancement factors as high as 7360

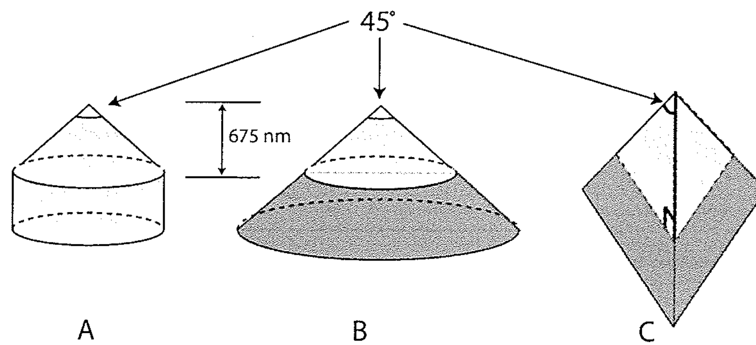


Figure 1.4: Various tip geometries modeled in [24]: a) quasi-Infinite extention metal conical tip b) metal/dielectrical conical tip c) right angle pyramid metal tip.

were calculated at the tip, using the pyramidal geometry in Figure 1.4c.

Apertureless NSOM apparatus

The apparatus used in a typical apertureless system is shown in Figure 1.5.

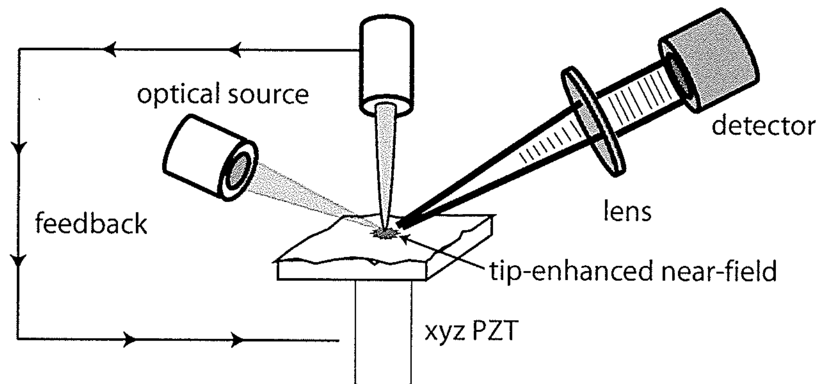


Figure 1.5: Basic apparatus use in apertureless NSOM.

The light source is focused on to the sharp probe tip. As we have seen, this leads to an enhanced near field. The interaction between the tip and the sample causes scattering that is detected in the far field. The tip/sample distance is independently controlled using force feedback. The signal/noise ratio can be improved by oscillating the tip and then extracting the detected field at the

oscillation frequency[25].

Advantages and Disadvantages of Apertureless NSOM

Because of the near-field amplification made possible by the probe tip, apertureless NSOM possess a much higher power near-field power density than ANSOM. This can allow for localized spectroscopy and material processing, which is a significant advantage of this technique. There are other significant advantages to apertureless NSOM. The probe tip's manufacture is greatly simplified, and near-field studies can even be accomplished with a traditional AFM probe[26]. The simpler probes give other advantages. Since the system no longer uses fiber optics, a wider range of wavelengths can be used. Additionally, without the aperture the probes can be much smaller. This reduces the effects of topography on the results[26].

Unfortunately, apertureless NSOM suffers from a significant limitation that has limited its deployment. The light that is directed to the tip also lands on the sample, creating a significant amount of background radiation. The interpretation of images is further complicated due to the presence of multiple scattering sources such as the tip dipole, its image in the sample and scatterers on the tip shaft[27]. Until the contrast mechanisms are fully understood, which are often dependent on the apparatus, only limited information can be gleaned from images taken using this technique, despite their high resolution.

1.2.3 Scanning Tunneling Optical Microscopy (STOM)

An interesting technique that has features of both apertured and apertureless NSOM is scanning tunneling optical microscopy (STOM) [28], also known as photon scanning tunneling microscopy (PSTM) [29]. The basic apparatus behind this technique (which is very similar to traditional STM), is shown in Figure 1.6. In this technique, a sharp, optical probe is scanned a very short distance (roughly 10 nm) from the surface. The sample is placed over a prism which is struck by a beam of light. The beam is in total internal reflection, which results in a evanescent field that interacts with the sample. This requires that the sample be very thin. This evanescent field is modulated by the sample, and the result collected by the sharpened probe. Because the probe

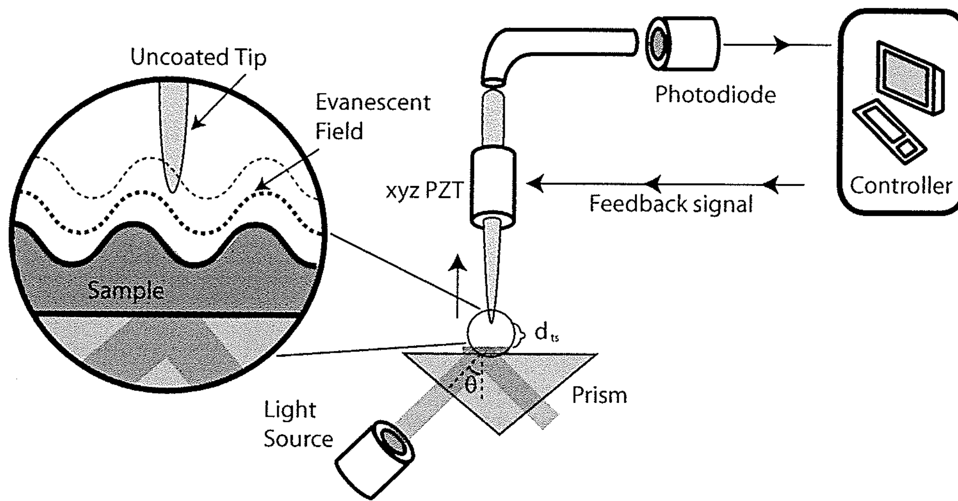


Figure 1.6: Basic apparatus for PSTM/STOM.

is used as a collector, there is no need to coat the tip with metal, and this has actually been shown to reduce the efficiency of the technique[29]. Through the use of shear-force distance regulation, lateral resolution below 10 nm has been achieved[30].

Like apertureless NSOM, the contrast mechanism in PSTM is rather complicated due to the numerous interactions taking place at the same time[30]. Since it does not offer any significant advantages over the other two NSOM techniques, PSTM is typically only used in certain niche applications.

1.3 Nanoantennas

The limitations of apertured and apertureless NSOM lead us to consider other arrangements for near-field sources. The ideal combination would be a source that combines the advantages of the apertureless and apertured configurations. It would offer a high energy density combined with a minimal spot size, have a similar contrast mechanism to optical microscopy and contribute minimal background radiation. Many of these advantages can be provided with a source that employs the basic ideas used in ANSOM, but replaces the metallized tapered fiber with an antenna structure. In ANSOM, the small aperture in the metal coating allows only a fraction of the incident light to reach the sample. With a properly designed antenna structure, the incident light that is

normally lost in the metal, either through reflection or absorption[5], could instead be channeled towards the illumination region. As wavelengths shorten, the optical properties of metals change so that they behave less like conductors and more like lossy dielectrics. An intuitive explanation of why this occurs is outlined in [31]. In a perfect metal the electrons are free and can respond instantly to any applied electric field, which means that $\epsilon = \pm\infty$. Since we say that there is zero electric field inside a perfectly conducting metal, the electrons respond by canceling the field, thus $\epsilon = -\infty$. Since the electrons cannot respond instantly, in actuality ϵ is a large negative number. As the frequency of the electric field increases, the electrons find it more difficult to respond, and this number gets progressively smaller. Additionally, there is an energy loss from resistive scattering, and this gives a positive imaginary component to ϵ. These parameters can be related to the complex index of refraction by the expression

$$n + ik = \sqrt{(\epsilon_r + i\epsilon_i)}. \tag{1.3}$$

The lossy dielectric behaviour of metals at high frequencies results in these small-scale antennas acting as nano-resonators. By properly tuning the resonance through the size of the antenna, an even higher proportion of the energy can be channeled towards the illumination region. Since the spot size is so small, only a very low illumination power would be required to achieve very high power densities. The very high amplification factor means that the background signal which was an issue for apertureless NSOM would be significantly reduced.

These near-field antennas are known as nanoantennas, since to achieve maximal resonance in the optical and near-infrared regimes they must be on the order of a few hundred nanometers in length. The bowtie antenna configuration is a natural choice for nanoantennas, as it offers good broadband characteristics as well as a relatively simple geometry. The first study outlining the use of nanoantennas to couple energy to the near-field was performed in [32]. As a proof of concept, the authors created a scale model operating at 2.2 GHz (λ = 13.6 cm). Figure 1.7 outlines the antenna geometry as well as the illumination and measurement apparatus. The total length of the antenna, L, was 36 cm (2.6 λ) and the width of the gap, d, was 1 cm, which was the same as the metal thickness. The antenna was excited by placing it 2.5 cm beyond the

open end of a rectangular waveguide. The polarization of the E field exiting the waveguide was in the same direction as the length of the antenna.

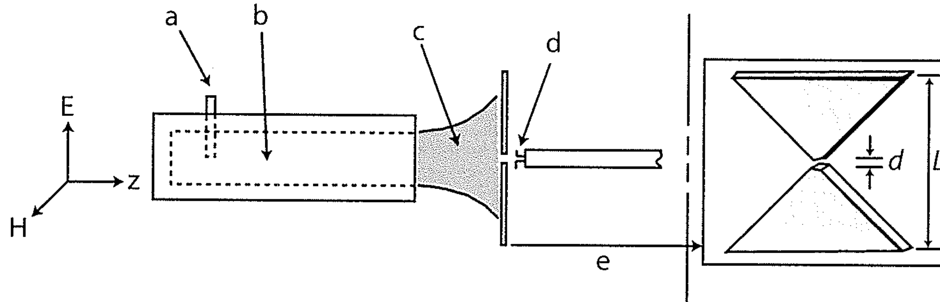


Figure 1.7: 3.2 GHz scale model illumination, measurement and antenna apparatus. Reused with permission from Robert D. Grober, Robert J. Schoelkopf, and Daniel E. Prober, Applied Physics Letters, 70, 1354 (1997). Copyright 1997, American Institute of Physics.

To measure the near-field of the scale model, a dipole probe was scanned over the antenna 0.5 cm (roughly $\lambda/30$) away. For comparison, an image was also taken a distance of 2.5 cm from the waveguide opening, without the antenna present. The results of these measurements are shown in Figure 1.8. In 1.8b, we can see how the antenna structure confines the energy to a region approximately $\lambda/10$. By comparing the ratio of the far-field power with and without the antenna in place, the coupling efficiency can be determined. For the setup in Figure 1.7, the efficiency was on the order of 2%. Higher transmission efficiencies were obtained by adjusting the distance between the antenna and the opening of the waveguide to create a resonant cavity between the two. By placing the antenna 6.8 cm ($\approx\lambda/2$) from the waveguide opening, transmission efficiencies on the order of 30% were obtained.

Another advantage of bowtie antennas is that they could be fabricated on pyramidal cantilevered tips. Figure 1.9 shows a maskless method[33] that could be used to fabricate this device. The authors constructed a scale model using this technique, operating at 3 GHz (scale factor 10^4). To analyze the performance, the probe was excited using a pyramidal horn with a C-band waveguide operating at 3.25 GHz, and the near-field was measured using a dipole probe connected to a power meter. Figure 1.10 shows the results of the near-field measurements.

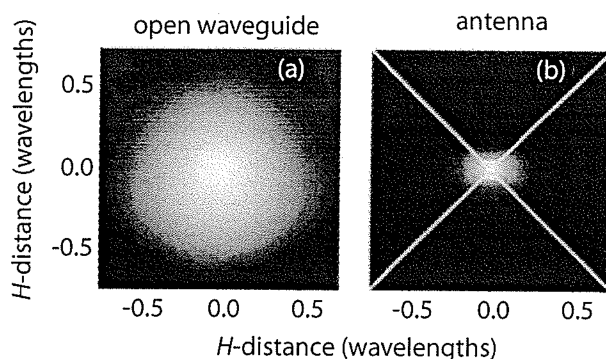


Figure 1.8: Field intensity 2.5 cm from waveguide opening (a) without antenna (b) with antenna. Reused with permission from Robert D. Grober, Robert J. Schoelkopf, and Daniel E. Prober, Applied Physics Letters, 70, 1354 (1997). Copyright 1997, American Institute of Physics.

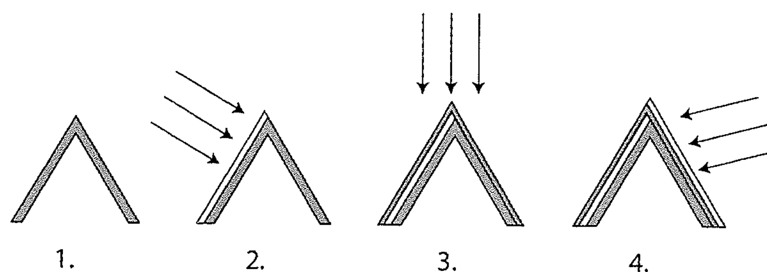


Figure 1.9: Maskless method for fabrication of bowtie antennas of a pyramidal cantilevered tip. (1) Begin with a standard dielectric tip, (2) sputter metal from one side, (3) coat a thin layer of tip material, (4) sputter metal to the opposite side. This technique will create the two halves of the bowtie on the sides of the tip, with the gap region defined by the thickness of the dielectric layer. The shadowing effect will leave the other sides uncoated by metal.

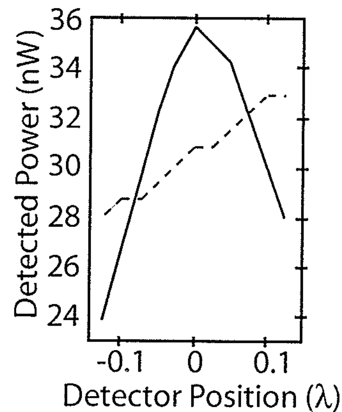


Figure 1.10: Near-field measurement of pyramidal antenna. Reproduction of Fig. 4 of [33].

Although the results of [32] and [33] clearly show that a bowtie geometry couples energy in to the near-field, the extension of the technique in to the near-infrared and optical regimes is not a trivial matter. Firstly, there are significant differences between the material properties of metals in the microwave and optical regimes that could change the device behaviour. To get an idea of the significance of this, the complex permitivity of Gold at 600 nm is roughly $\epsilon = -8.26 + i1.12$[34]. The longest wavelength with a listed permitivity value in [34] is 10 μm, and already the permitivity is $-2842 + i1339$. At 3 Ghz the wavelength is 10 mm, and the permitivity will be so negative that it can be approximated as $-\infty$.

In addition to the material difference, the small scale required for optical resonance means that fabrication of the antennas poses a significant challenge. Finally, the near-field at optical wavelengths will only extend a distance of a few tens of nanometers, making the measurement process exceedingly difficult.

The first characterization of the near-field coupling properties of nanoantennas in the visible regime was reported in early 2005[3]. The antennas were fabricated via electron beam lithography of gold on a fused silica cover slip, as described in [35]. Each half of the bowtie was approximately 75 nm long, with a material thickness of roughly 18 nm. The antennas were excited using a mode-locked Ti:sapphire laser operating at 830 nm, which is the wavelength at which the antennas should be resonant. The strong local fields of the antennas

leads to detectable two photon absorption[3]. This excites electrons from the *d* valence band to the *sp* conduction band, which causes the phenomena known as two-photon excited photoluminescence (TPPL). TPPL has a nonlinear dependance on the excitation energy, and can be used to determine the intensity of the field in the metal of the antenna[3]. Figure 1.13 shows the TPPL excitation measured from the antennas.

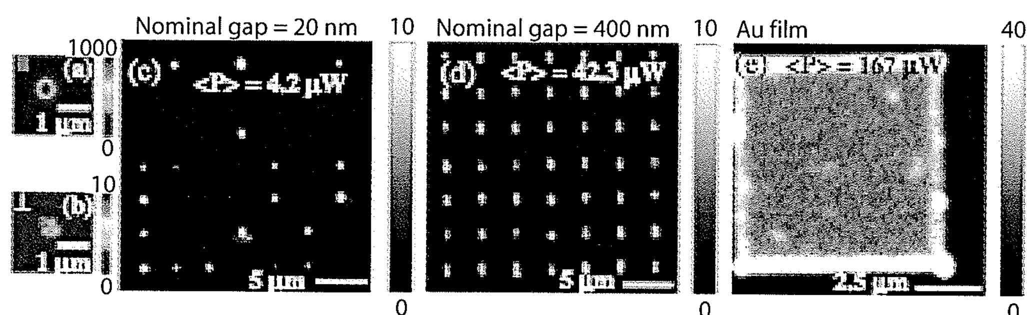


Figure 1.11: TPPL excitation of bowtie array (a) and (b) showing the dependance on the polarization of the incident light. (c) and (d) show the luminescence of bowtie arrays with nominal gap sizes of 20 nm and 400 nm, respectively. (e) showns the luminescence from a smooth Au film. Reprinted figure with permission from [3]. Copyright 2005 by the American Physical Society.

A comparison of the predicted and measured field enhancement versus gap width is given in Figure 1.12. The simulation results agree well with the measurements for gap sizes below 30 nm, but between 40-60 nm the results begin to diverge. The results near 150 nm and 400 nm seem to be in good agreement with the simulation results. [3] states that this is likely due to the fact that the excitation condition in the FDTD simulation closely approximations the experimental condition for small gap bowties, but when experimenting on medium sized gaps the triangles are no longer uniformly pumped by the diffraction limited focused laser spot. For the 400 nm gap, when the TPPL data was gathered, each half of the antenna was stimulated independently and the results added, since the antenna was much larger than the spot size. This geometry was also too large to simulate, so a single triangle was simulated and the results doubled.

Shortly after the measurement of the fields in the metal, the field enhancement in the gap region was measured [4]. The authors fabricated strips with a small gap at the center, and illuminated the antennas with a picosecond laser

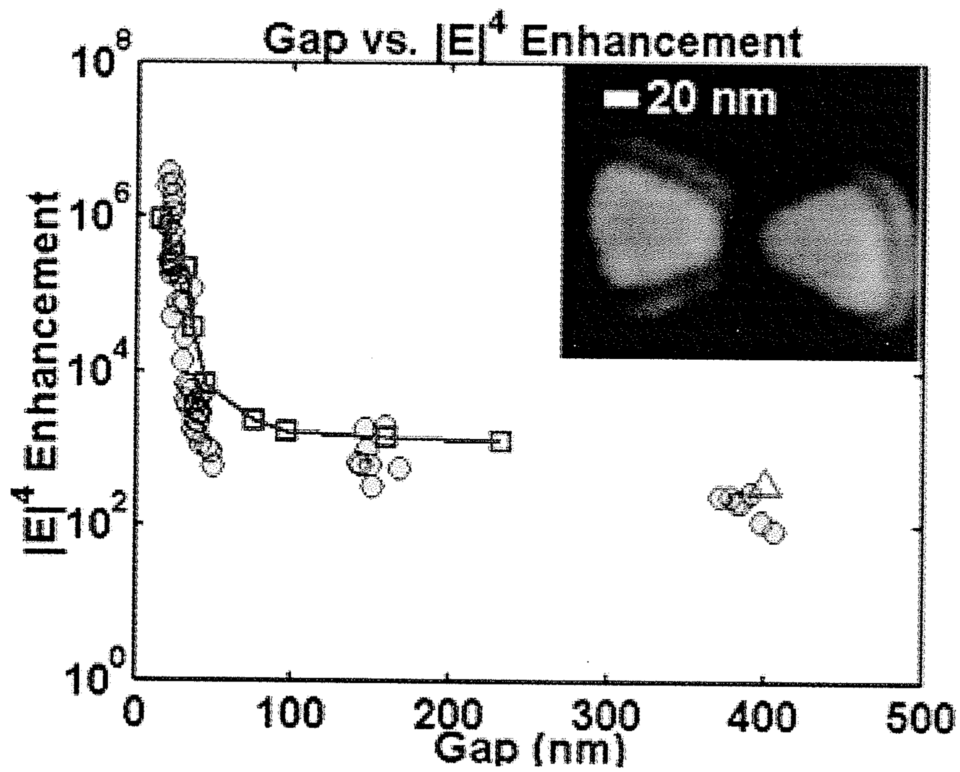


Figure 1.12: Comparison of theoretical versus measured field enhancement versus gap width. Inset shows an SEM image of an antenna with a 22 nm gap. Reprinted figure with permission from [3]. Copyright 2005 by the American Physical Society.

pulse. In addition to TPPL generation, the authors measure a fourth order optical nonlinearilty known as White-Light Super Continuum (WLSC)[36][37]. This effect occurs in some dielectric materials such as glass, water, and air, but not in gold. Thus by measurement of the WLSC generation one can characterize the field enhancement in the gap region without interference from the field in the metal. The mechanism behind WLSC is not well known[4], but appears to require a pulse length in the picosecond range. WLSC measurements were taken following illumination using an inverted confocal microscope. The results of these measurements for vertical and horizontal polarization with different antenna orientations are shown in Figure 1.13.

Figure 1.14 shows the measured WLC spectra for an antenna. For lower powers the spectra decreases linearly as the wavelength decreases, which is typical for TPPL of gold. At higher powers the peak at 560 nm begins to dominate, which is characteristic of WLSC. Figure 1.15 shows the WLC power dependence on the magnitude of the incident laser pulse. Figure 1.16 shows the both the measured and simulated behaviour of the WLC power versus the antenna length. In both cases we can see the existence of a resonant peak.

[3] and [4] both use ultrafast lasers and nonlinear effects for characterization. Recently [2] reported characterization of a bowtie shaped aperture illuminated with an standard Argon laser with λ=458 nm. The bowtie aperture, shown in Figure 1.17 along with other aperture types, has the advantage of blocking all of the stray light. The measurement was accomplished using an apertured NSOM apparatus with the cantilevered probe operating in collection mode. This system has a sub-100 nm resolution. An image taken with this system showing the near field coupling is shown in Figure 1.18. Using photon counts received from the NSOM probe, the authors estimate the antenna efficiency to be near unity over the probe opening area.

1.4 Outline

This work attempts to give further insight in to the behaviour of nanoantennas, as well as offer an alternative method to their characterization. The previous studies on nanoantennas have been limited to analysis of their operation in the visible regime. In this work we explore nanoantennas operating at 1550 nm.

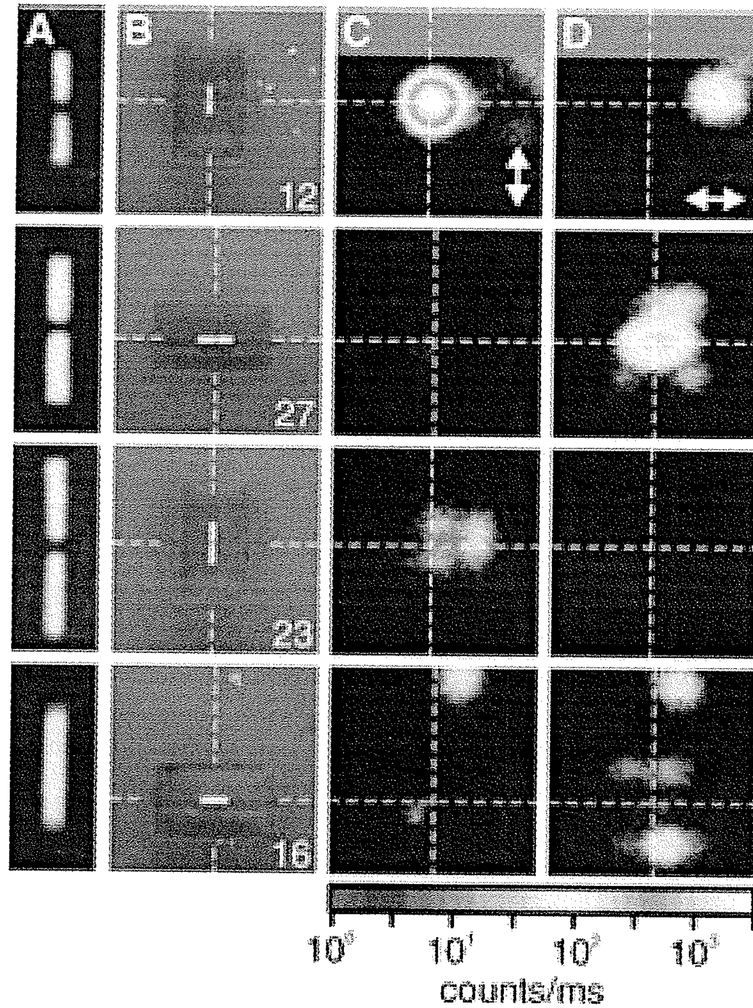


Figure 1.13: Nanoantennas and WLSC generation. A) and B) SEM images C) and D) WLSC generated from vertical and horizontally polarized laser pulses. A) is 450 x 180 nm^2, B), C) and D) are 2 x 2 μm^2 . From [4]. Reprinted with permission from AAAS.

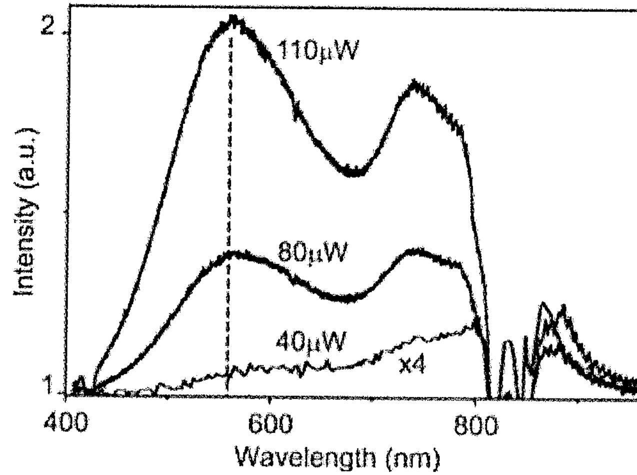


Figure 1.14: Measured WLC spectra for different excitation powers. From [4]. Reprinted with permission from AAAS.

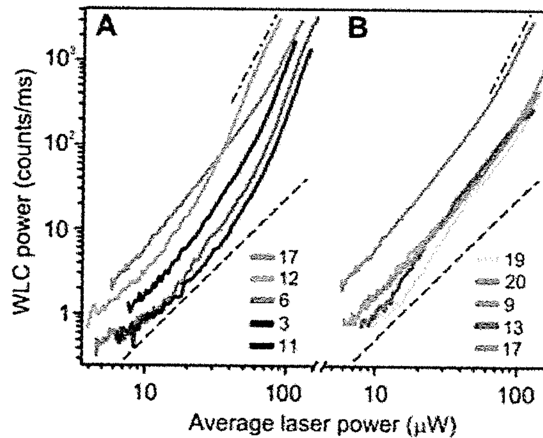


Figure 1.15: WLC power versus average incident laser power for different antenna lengths, sorted by dominant nonlinearity order. Dashed lines: (a) fourth order, (b) second order. From [4]. Reprinted with permission from AAAS.

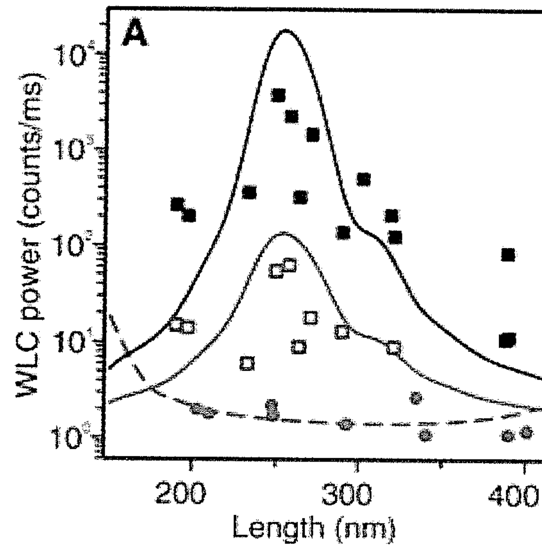


Figure 1.16: Simulated and measured relationship of WLC power versus antenna length. Filled and empty squares: Measured power at 110 μW and 30 μW, repectively. Solid red (bottom) and black (top) lines: $R^4(L)$ and $R^2(L)$ factors for antennas. Dashed red line: $R^2(L)$ for metal stripe. From [4]. Reprinted with permission from AAAS.

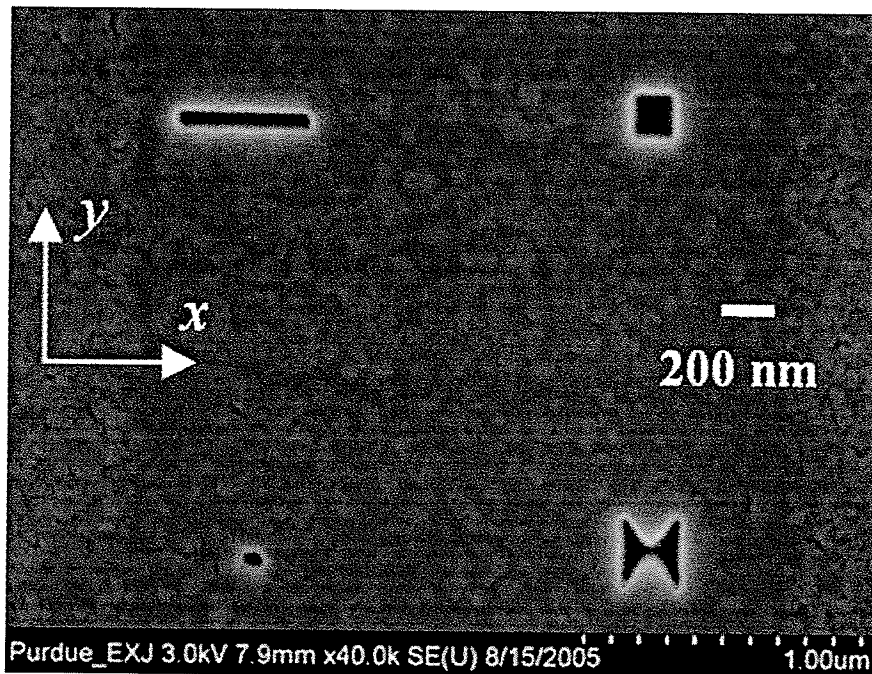


Figure 1.17: SEM image of aperture array. Reused with permission from [2]. Copyright 2006, American Institute of Physics.

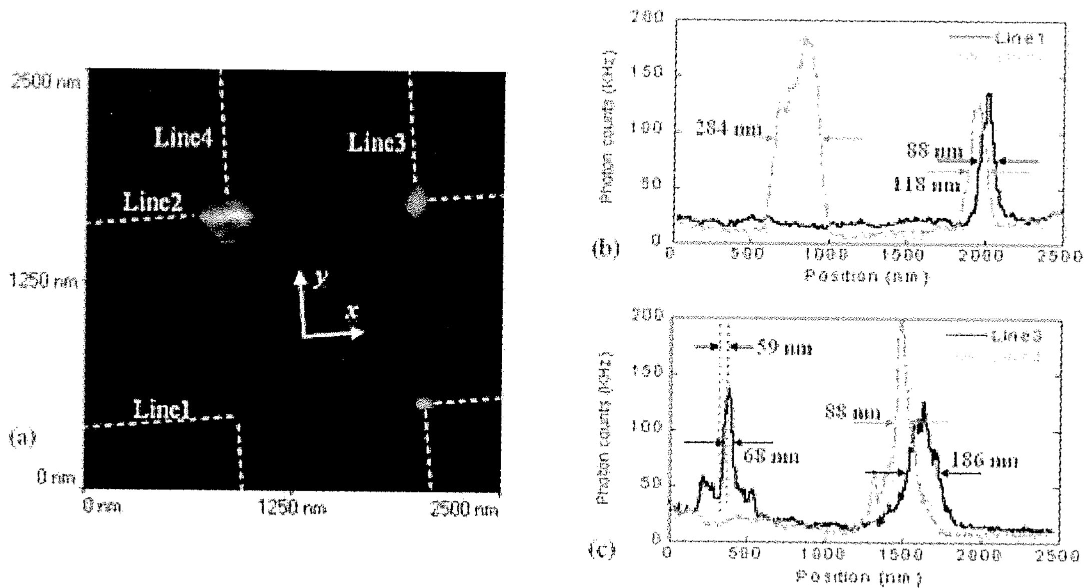


Figure 1.18: NSOM image of aperture array. (a) Near-field of (clockwise from top left) rectangular aperture, large square aperture, bowtie aperture, small square aperture. (b) Photon count along lines 1 and 2. (c) Photon count along lines 3 and 4. Reused with permission from [2]. Copyright 2006, American Institute of Physics.

We begin by simulating the behaviour of these devices at 1550 nm and then examine the effects of parameters such as material selection, antenna length, antenna thickness, and the spacing between the probe halves. An apparatus for exciting the antennas inside a scanning probe microscope (SPM) is designed. Following this, two methods of fabrication are explored: fabrication through focused ion beam milling, and through the use of electron beam lithography. Finally, we explore two techniques for the characterization of these antennas. The methods outlined in Chapter 5 have the advantage of being broadband techniques, as well as being independent of the illumination source. In the other published techniques for characterizing nanoantennas the illumination source is a key part of the experiment[3][4], or the measurement technique is restricted to a particular spectra at a time[2]. Finally, our characterization methods should yield a better picture of the spatial distribution of the near field.

Chapter 2

Simulation

2.1 Introduction

Our first task is to determine the optimal geometry for our antennas. For this, we turn to electromagnetic simulations of the antenna to determine the near-field intensity. For this task we have chosen to make use of the High-Frequency Structure Simulator (HFSS) package from the Ansoft Corporation[1]. While primarily used for RF simulations, the package can be used to simulate EM effects at optical frequencies as well. From these results we can determine the requirements of a potential fabrication process, and begin our manufacturing stage. Note that all of the optical property data in this chapter is taken from [34].

2.2 Material Properties

The main difference between simulations at optical frequencies and simulations at lower frequencies is the significant change in the physical properties of metals. We can easily extract this information from the optical properties. Beginning with the values of n and k listed in [34], we use the relation

$$\tilde{\epsilon} = (n + ik)^2. \tag{2.1}$$

[1] http://ansoft.com/

This means that the real part of the permittivity, ϵ_r, is

$$\epsilon_r = n^2 - k^2, \tag{2.2}$$

and the imaginary part, ϵ_i, is

$$\epsilon_i = 2 \times n \times k. \tag{2.3}$$

The ratio of ϵ_i/ϵ_r is known as the *dielectric loss tangent*, and describes how lossy the material is. In considering metals for use in our fabrication, we check for those which have a low loss tangent in the wavelength range that we are interested in. Table 2.1 shows the loss tangent of several metals at 1550 nm and 830 nm.

Table 2.1: Dielectric Loss Tangent of several metals at 1550 nm and 830 nm.

Material	1550 nm Loss Tangent	830 nm Loss Tangent
Aluminum (Al)	-0.1814	-0.740
Chromium (Cr)	-6.213	-71.66
Copper (Cu)	-0.1475	-0.09910
Gold (Au)	-0.1143	-0.06984
Iridium (Ir)	-0.8413	-1.297
Iron (Fe)	-1.921	-3.429
Molybdenum (Mo)	-0.4531	14.83
Nickel (Ni)	-1.314	-1.666
Platinum (Pt)	-3.500	-1.724
Silver (Ag)	-0.1009	-0.05276
Tungsten (W)	-1.158	4.488

If we wish to be able to simply extend our results to the visible spectra, suitable metals are gold, silver, aluminum and copper. Plots of the behaviour of the real and imaginary parts of their dielectric properties are given in Figures 2.1 and 2.2. Note that the behaviour in Figure 2.1 is to increase with frequency, as explained in Section ??. The loss tangent across the same range is given in Figure 2.3.

Once we have the electrical properties, the following steps are used to create parametric materials in HFSS:

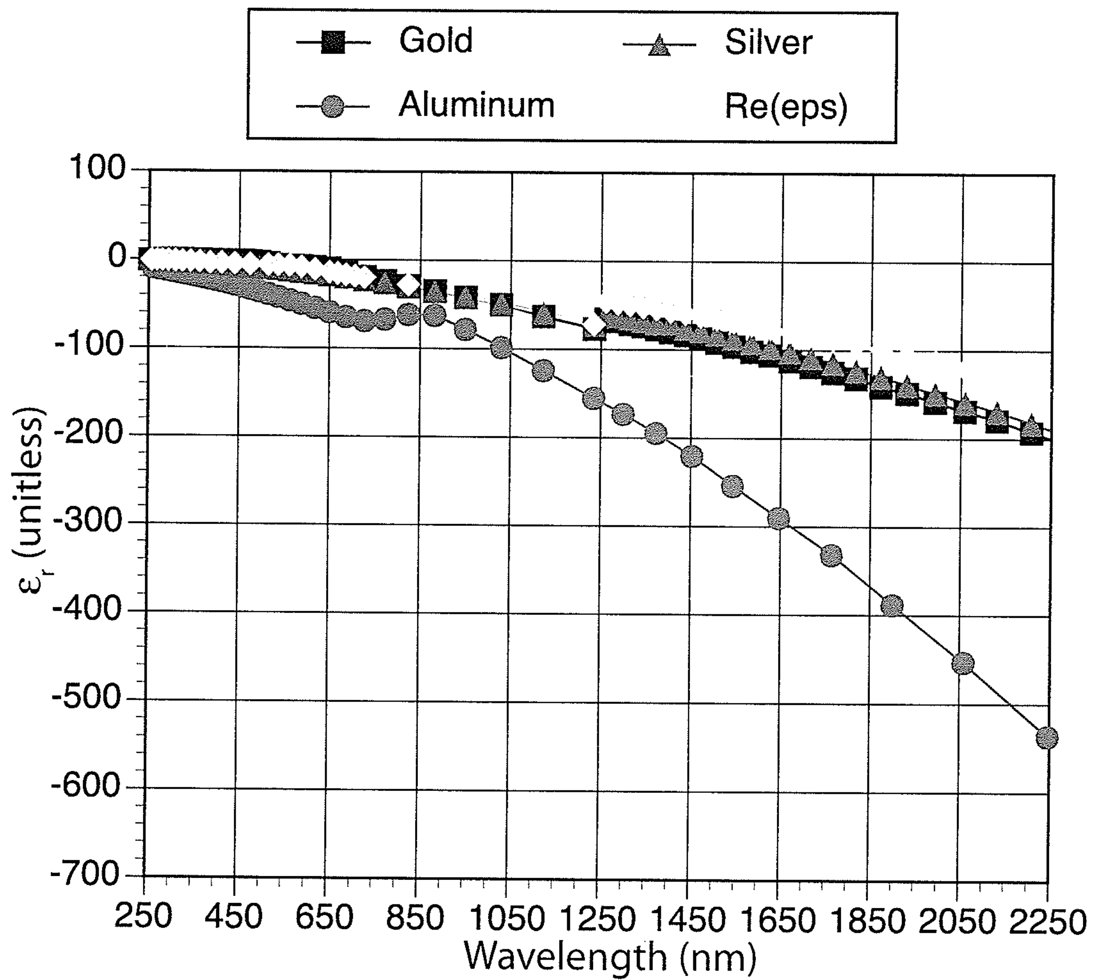


Figure 2.1: Relationship of ϵ_r versus wavelength.

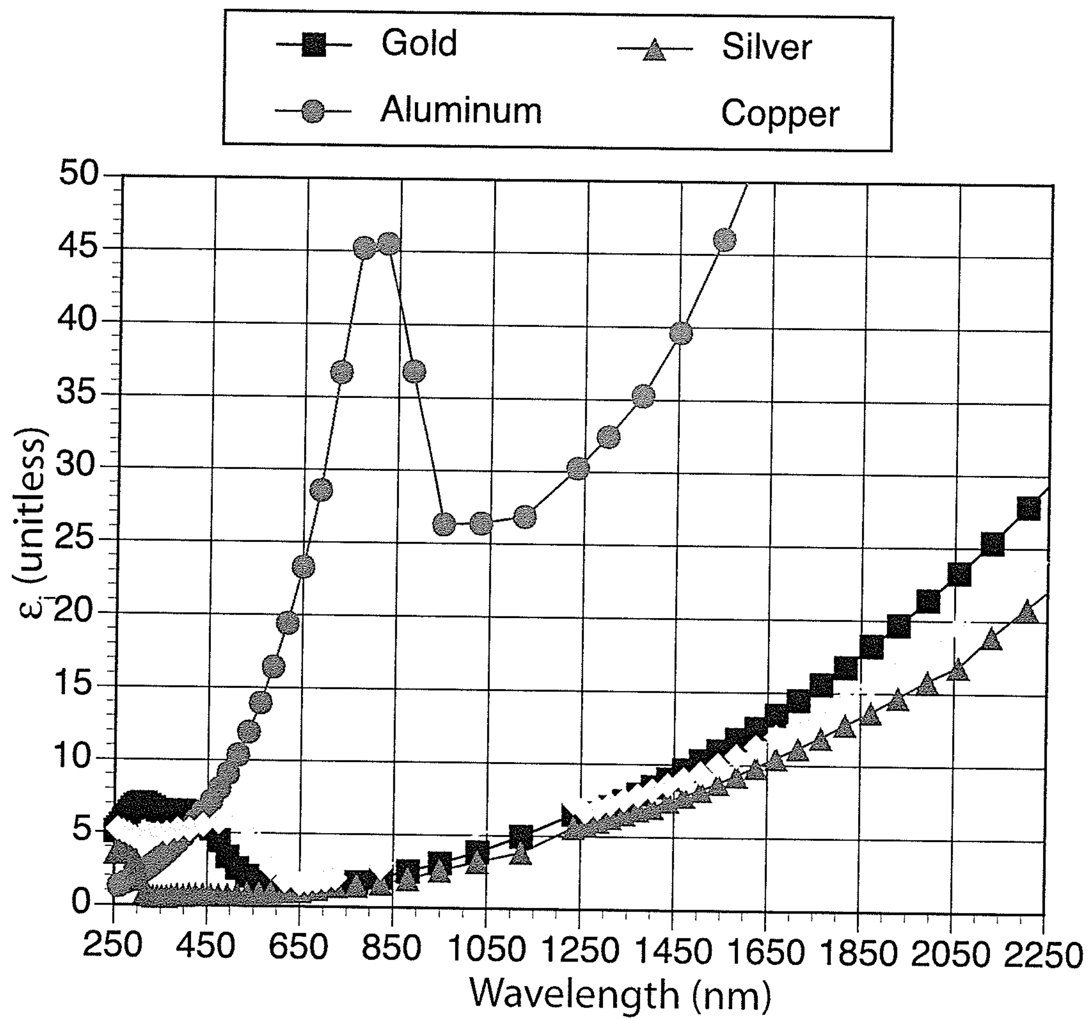


Figure 2.2: Relationship of ϵ_i versus wavelength.

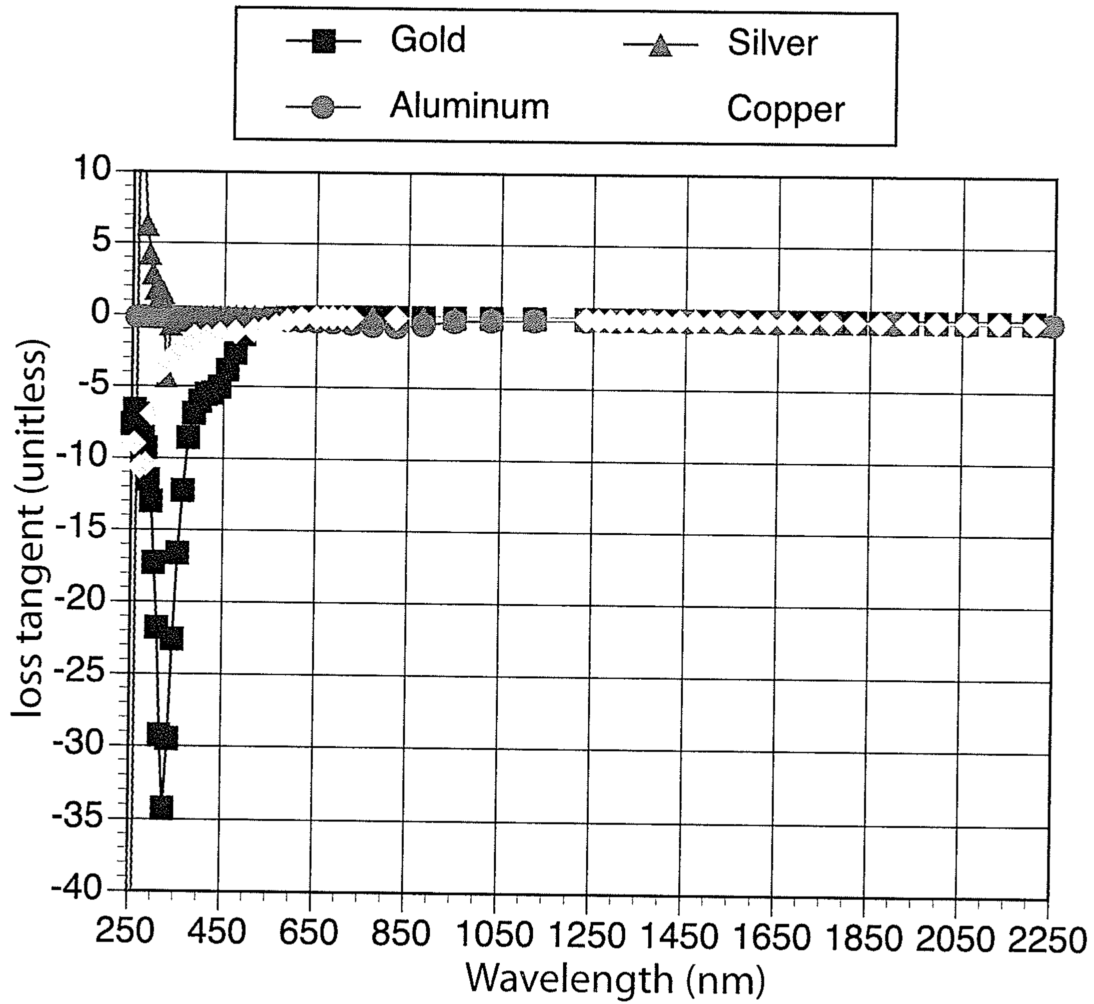


Figure 2.3: Relationship of the dielectric loss tangent (ϵ_i/ϵ_r) versus wavelength.

1. We define a new data set for the real part of the permittivity, ϵ_r and the dielectric loss tangent, ϵ_i/ϵ_r using `Project→Datasets`, and clicking the `Add` button. The data set can be either typed in, or imported from a tab delimited text file with the extention *.pts*. If a file is imported, HFSS will display the error *Non numeric data "" found in file*, but import the data properly regardless.

2. Define the new material with `3D Modeler→Assign Material` and clicking the `Add Material` button. Change the *Relative Permittivity* value to `pwlx(dataset(`ϵ_r *dataset name*`),Freq)` and the *Dielectric loss tangent* value to `pwlx(dataset(`*loss tangent dataset name*`),Freq)`.

2.3 Simulation Geometry

In order to create the element of our simulation, we employ a parametrically defined geometry. The parameters we will wish to vary are

- Antenna size *scale*: The total length of the metal in the antenna.
- Gap size *gap*: The minimum distance between the two metal halves.
- Antenna Thickness *height*
- Flare Angle *angle*: The total angle between the edges of a triangle with the same length as the antenna half, coming to a point at the center.

In order to create an antenna geometry based around these parameters, we define several other parameters:

- Design wavelength λ_d:
- Maximum x coordinate of the antenna max_x: $\lambda_d \times size/2$
- Maximum y coordinate of the antenna max_y: $max_x \times \tan(angle/2)$

All of these parameters are entered into the simulation by selecting the design and using the command `HFSS→Design Properties`.

The bowtie is created using the following process:

1. We create a circle with a 15 nm radius at the point (*gap*/2+15 nm, 0, 0) using the `Draw→Circle` command. The purpose of the circle is to blunt the tips of the bowtie, this simulating a geometry closer to one that would result from fabrication.

2. We then create a 4-sided polygon using the `Draw→Line` command. The vertices are at (*gap*/2+ 9.1 nm, 14.1 nm, 0)→(max_x, max_y, 0)→(max_x, -max_y, 0)→(*gap*/2+ 9.1 nm, -14.1 nm, 0)→(*gap*+ 9.1 nm, 14.1 nm, 0).

3. We combine the circle and the polygon by selecting both of them and using the `3D Modeler→Boolean→Unite` command.

4. The antenna half is given thickness using the command `Draw → Sweep → Along Vector`. The range of the sweep is set to (0, 0, 0)→(0,0, *height*).

5. We assign the a material type to the probe using `3D Modeler→Assign Material`.

6. The other half of the antenna is created using a mirror across the x-axis with the command `Edit → Duplicate → Mirror`. The mirroring vector is (0, 0, 0)→(1, 0, 0).

To give our simulation a boundary, we create an airbox around the antenna and assign it the task of absorbing the radiation. This is done as follows:

1. We create the air box shape using `Draw→Box`. The starting point is (-max_x - 1000 nm, -max_y - 1000 nm, -300 nm). The sizes are set to $2\times(max_x$ + 1000 nm), $2\times(max_y$ + 1000 nm) and 800 nm in the x, y and z directions respectively.

2. We set the material type inside the airbox to *Vacuum* using the command `3D Modeler→Assign Material`.

3. We then set the boundaries of the air box to absorb radiation using the command `HFSS→Boundaries→Assign→Radiation`.

Finally, the illumination of the antenna via an incident plane wave is done using the command `HFSS→Excitations→Assign→Incident Wave`. The vector input format is set to `Cartesian`, and the *k*-vector to (0,0,1).

2.4 Results

2.4.1 Basic Simulation

For our initial simulation, we began with a probe whose halves were each the same length as one half the incident wavelength of 1550 nm. The metal thickness was chosen to be 40 nm, the flare angle was chosen to be 40°, and the spacing between the tips was chosen to be 10 nm. This simulation was set to terminate following 10 passes, or when the difference in the average electric field between two intervals falls below 0.001 V/m, whichever occurs first. The convergence data for this simulation is given in Figure 2.4, showing that the simulation appears to be converging. Figure 2.5 shows an image of the

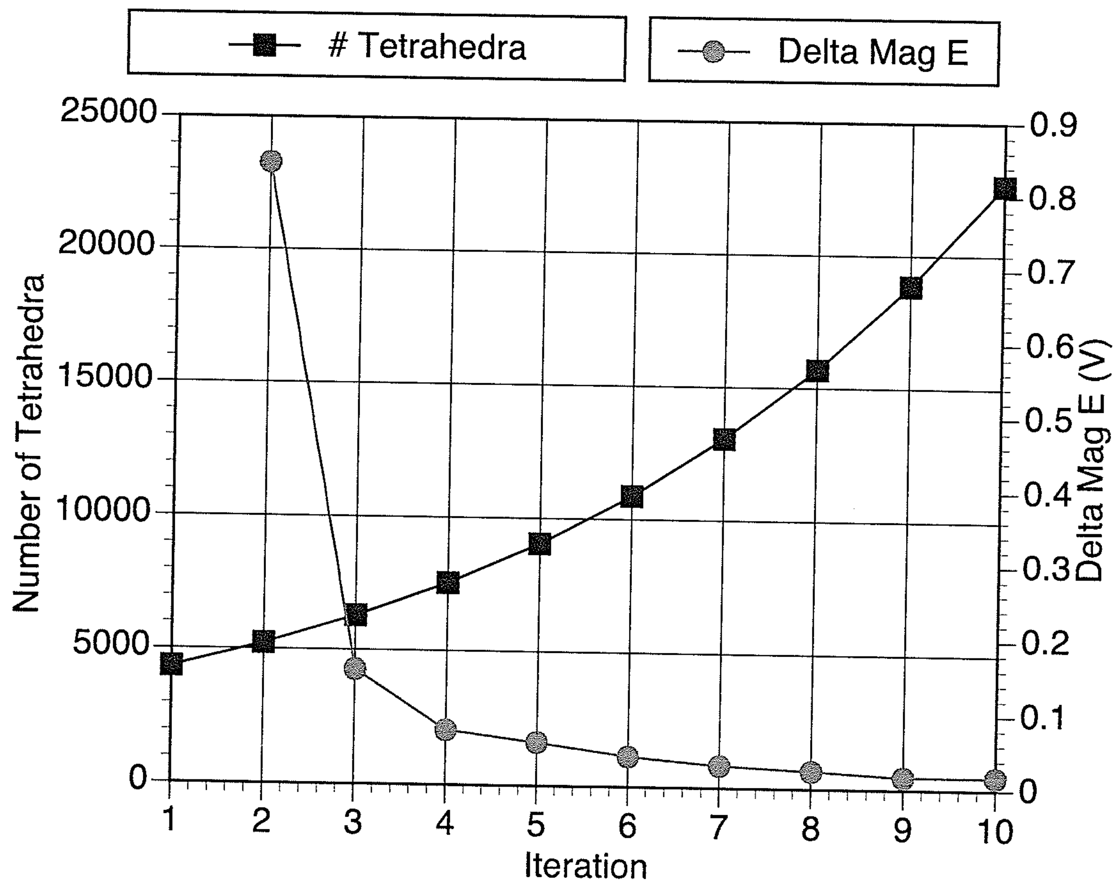


Figure 2.4: Simulation convergence data.

isovalues within the metal halves of the antenna. We can see that we expect a higher field intensity at the edges of the antenna, with the highest intensity

near the gap region. Figure 2.6 and 2.7 show the field along the x- and y-axes, 1 nm above the metal surface, for both x- and y-polarized incident radiation. Figure 2.8 shows the relationship between the electric field and the distance from the top of the antenna, again for x- and y-polarized radiation. Since the near-field enhancement occurs solely with and x-polarized illumination, we will restrict our remaining simulations to this type of excitation.

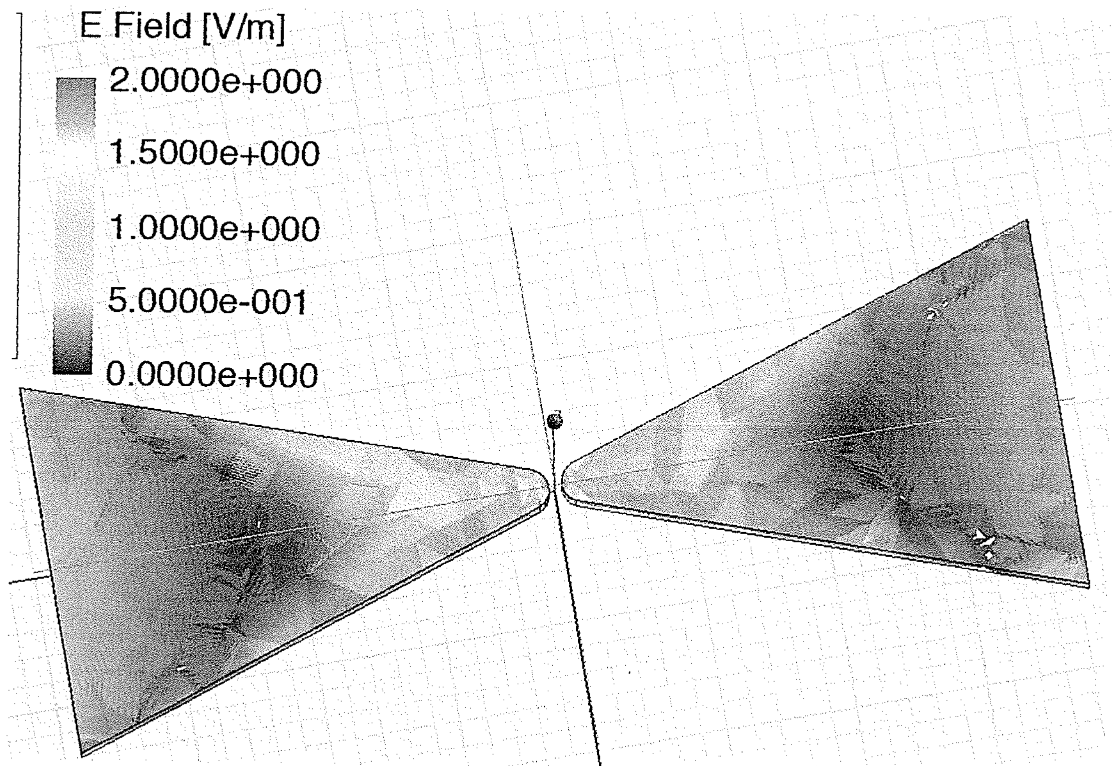


Figure 2.5: Electric field inside the metal arms of the antenna. λ = 1550 nm, Total Metal Length = 1λ, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

Figure 2.9 shows the electric field in the plane above the antenna, offering more evidence of electric field confinement and enhancement. Figure 2.10 shows the magnetic field, and we can see that a similar effect does not occur. If the gap is shorted current can freely flow between the two halves and the confinement occurs with the magnetic field rather than the electric field. This is illustrated in Figures 2.11 and 2.12.

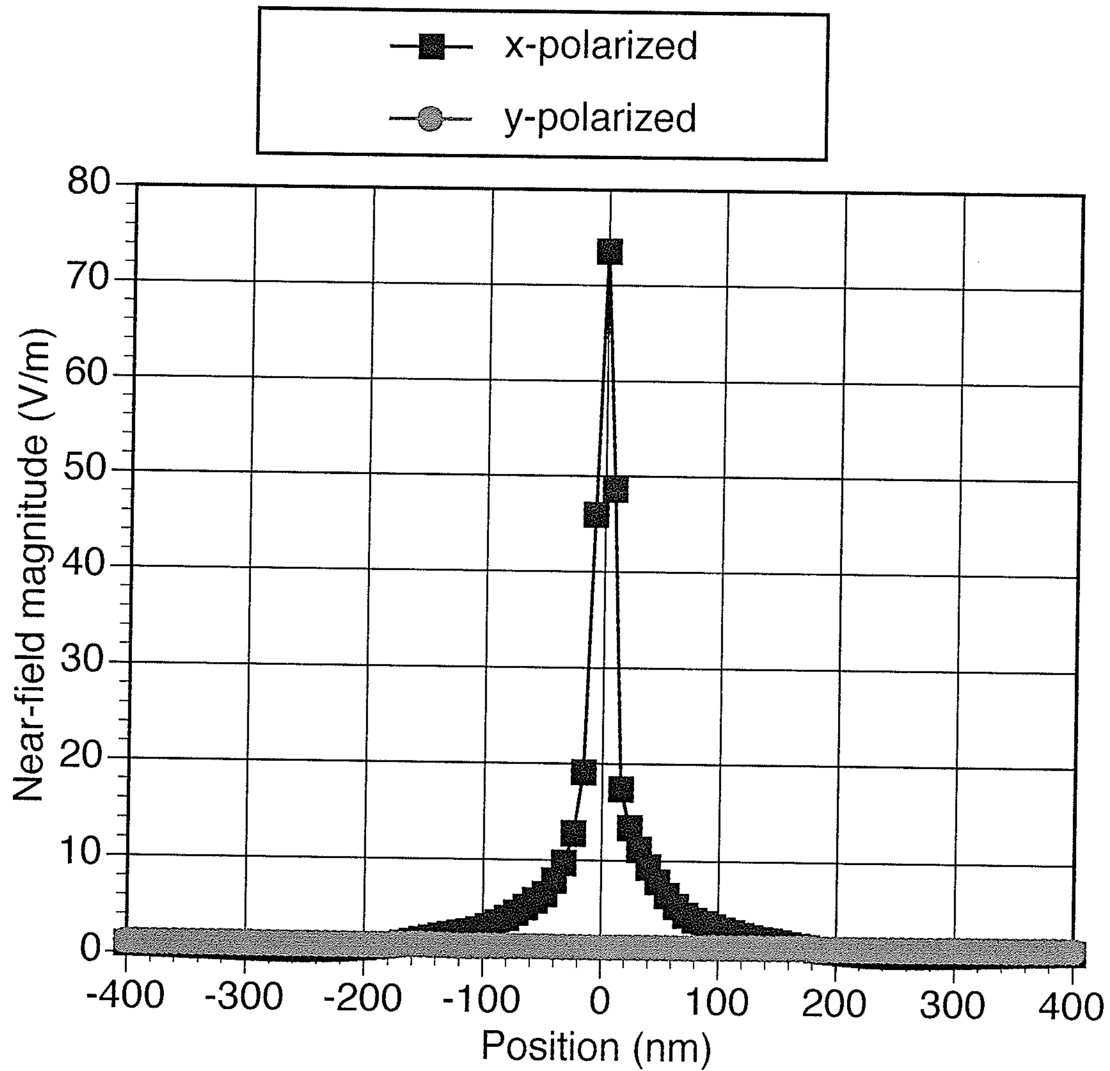


Figure 2.6: Electric-field along the x-axis at the top of the antenna 1 nm from the metal, given a 1 V/m linearly polarize plane wave excitation. $\lambda = 1550$ nm, Total Metal Length = 1λ, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

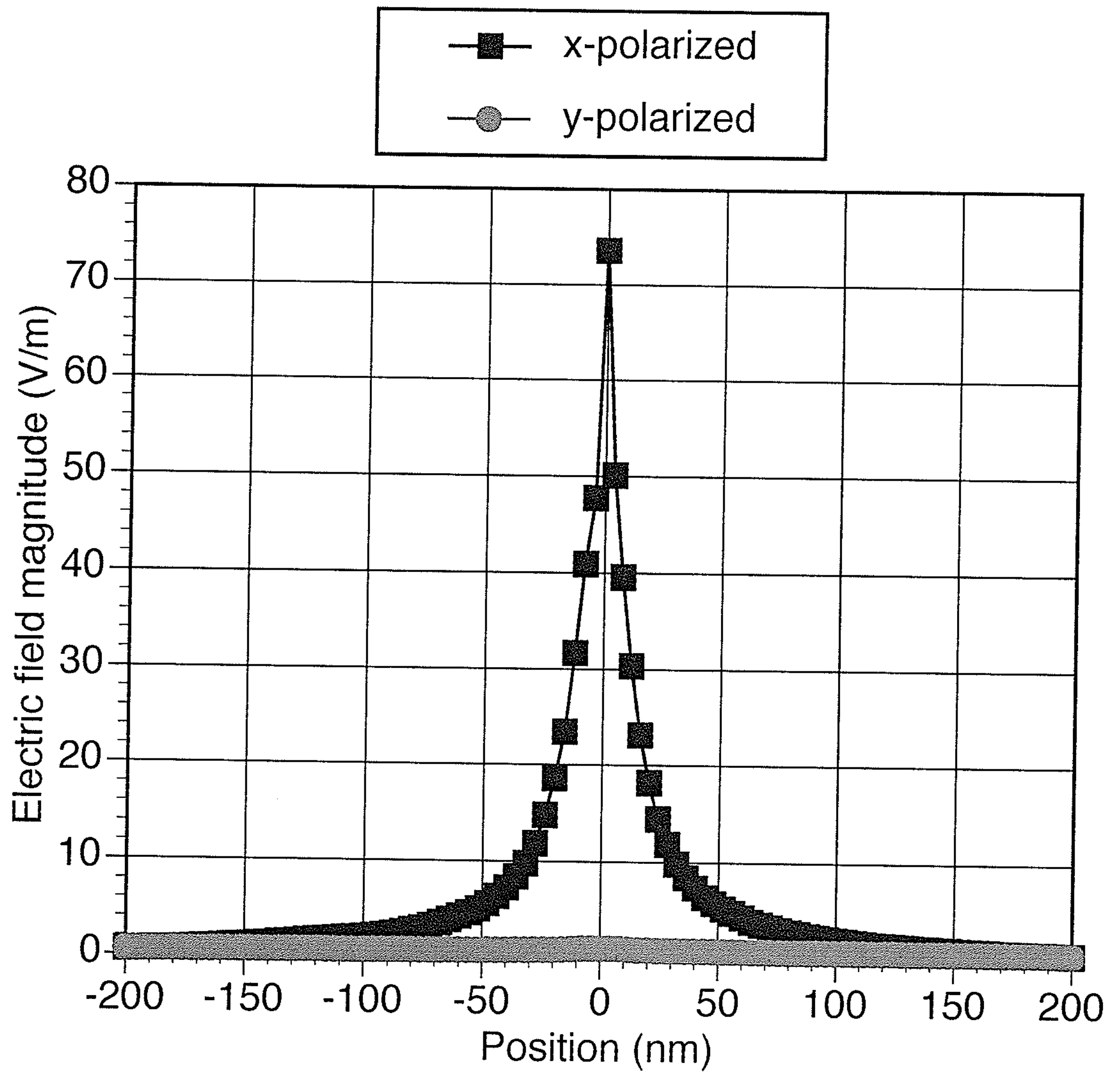


Figure 2.7: Electric-field intensity along the y-axis at the top of the antenna, 1 nm from the metal. λ = 1550 nm, Total Metal Length = 1λ, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

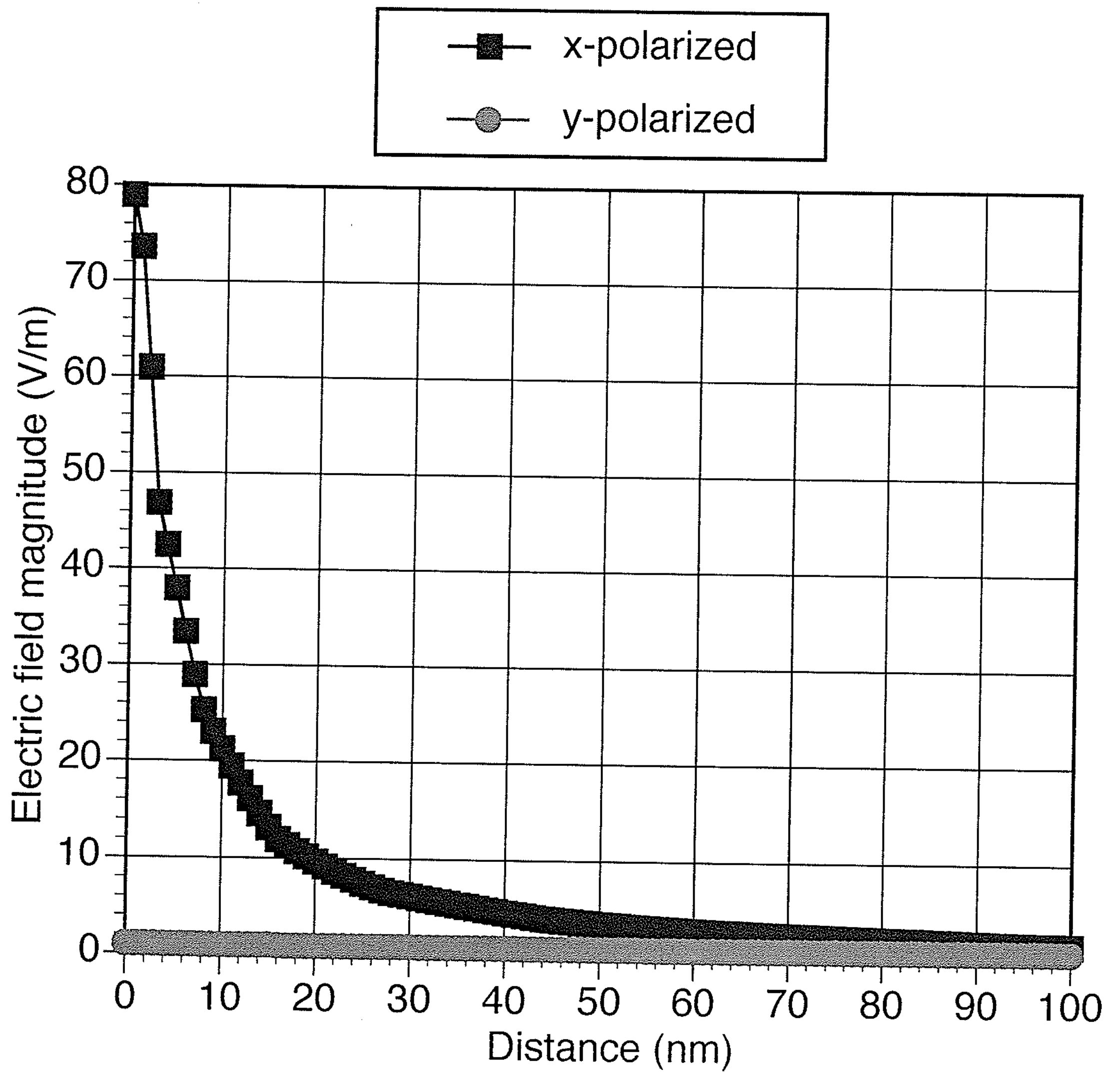


Figure 2.8: Electric-field versus distance away from antenna, given a 1 V/m linearly polarize plane wave excitation. $\lambda = 1550$ nm, Total Metal Length = 1λ, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

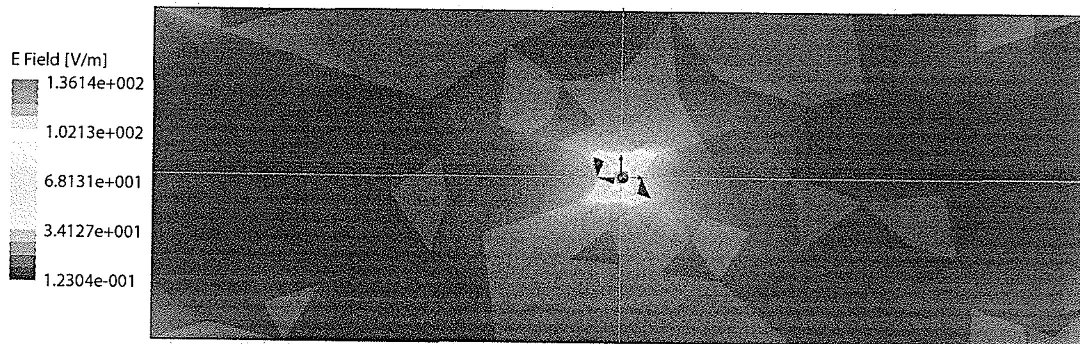


Figure 2.9: Electric field in the plane above the antenna. λ = 1550 nm, Total Metal Length = 1λ, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

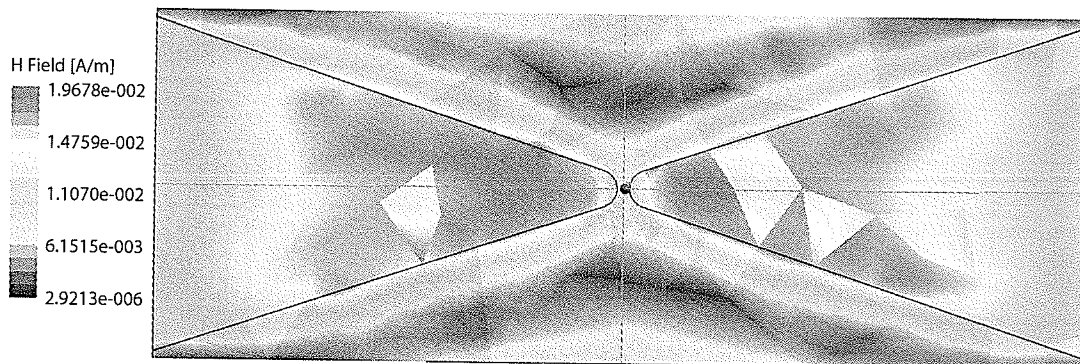


Figure 2.10: Magnetic field in the plane above the antenna. λ = 1550 nm, Total Metal Length = 1λ, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

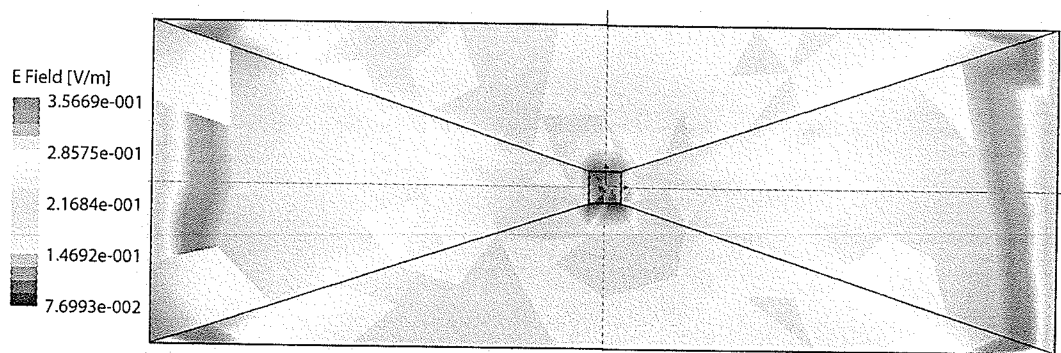


Figure 2.11: Electric field in the plane above an antenna with connected halves. λ = 1550 nm, Total Metal Length = 1λ, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

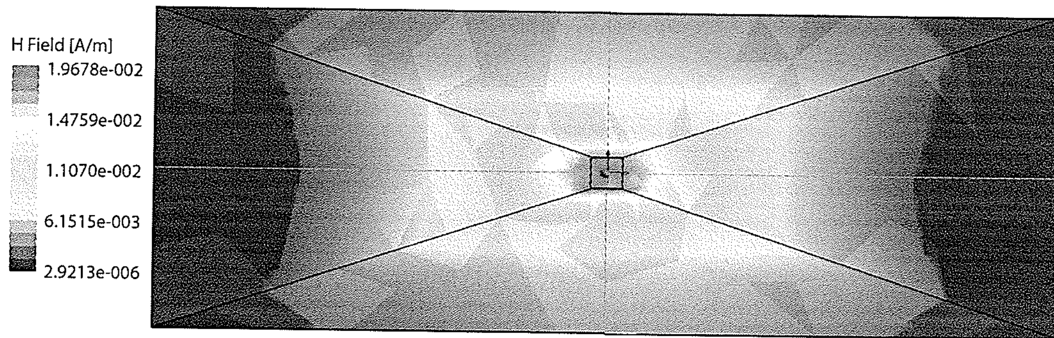


Figure 2.12: Magnetic field in the plane above an antenna with connected halves. λ = 1550 nm, Total Metal Length = 1λ, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

2.4.2 Parameter Variation

We next simulate the effects of modifying the various attributes of the antenna. This will allow us to establish an optimal design that is realizable with our chosen fabrication method. For these simulations we have used the same geometry as in the previous section, with the exception that one of the parameters is varied across a range of values. The amplification factor is defined as the electric-field intensity at the very center of the top of the antenna (i.e. the magnitude of the electric field at point (0,0,*height*)) when the antenna is excited by a 1 V/m x-polarized plane wave. Since parametric simulation are a time-intensive task, six iteration was deemed sufficient. The first parameter that is varied is one of the most critical to the antenna's functionality. The total length of the antenna determines whether the antenna is resonant with the incident electric field. This proves to be very important, and the maximal resonance occurs when $L \approx 0.5\lambda$. The results for varying the gap size of a maximally resonant probe are given below in Figure 2.14. This parameter also proves to be very important to the efficiency of the antenna. The two remaining parameters are less important to the antenna behaviour. Figure 2.15 shows the dependence of the antenna gain on the thickness of the metal layer. While this is not a significant effect, we can see that thinner metal is preferable, if thicker than 20 nm. The effect of the flare angle is shown in Figure 2.16. The effect is roughly linear when the antenna is maximally resonant with the incident field, but when it is not the effect on the flare angle is unpredictable.

Finally, one of the aspects of our antennas that we are interested in is how

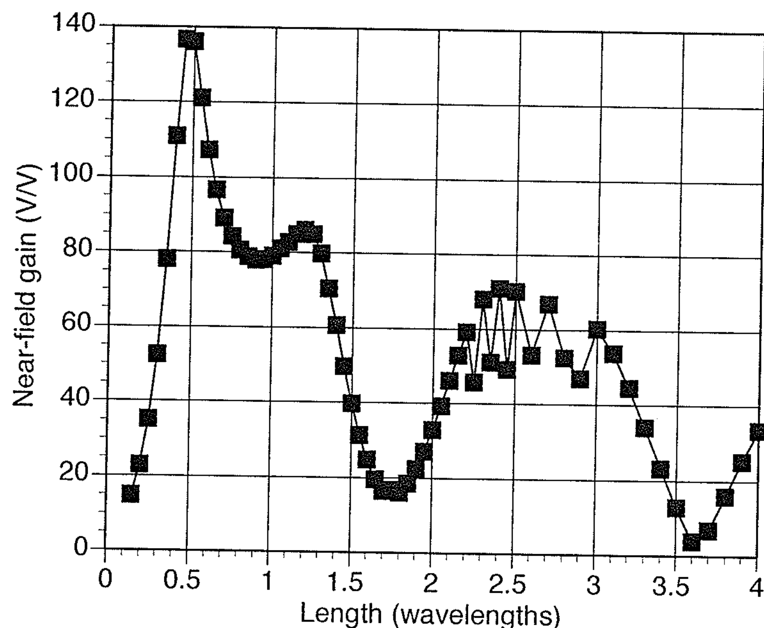


Figure 2.13: Amplification factor versus antenna length. λ = 1550 nm, Spacing = 10 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

it behaves across the electromagnetic spectrum. Figure 2.17 shows the effect of varying the incident wavelength on the antenna gain for our chosen antenna, assuming a gap size of 10 nm.

2.5 Conclusions

From the results of our parametric simulations, we now know the shape of an effective nanoantenna. Of the variable parameters, the metal thickness, antenna size and flare angle can all be controlled relatively easily. The antenna length was chosen to be the maximally resonant length, 775 nm. Because very small thicknesses led to reduced performance, and to allow for some error in the simulation and in fabrication tolerances, the metal thickness was chosen to be 40 nm. The flare angle was chosen to be 40°. Since it was assumed that the most difficult step in the fabrication process would be in creating the gap between the antenna, this requirement was merely stated to be as small a gap as was possible with the chosen technique.

We can also deduce that any analysis technique that is to characterize these antennas should be capable of examining the fields within approximately 30

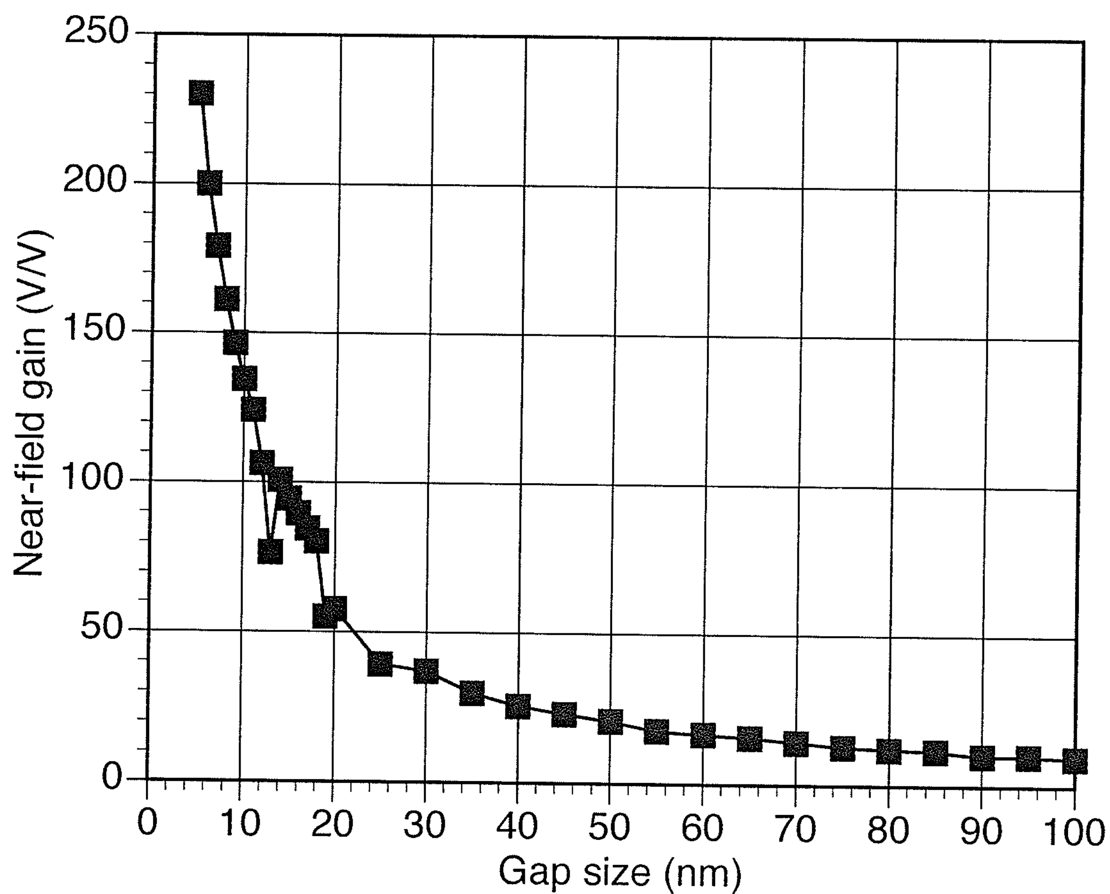


Figure 2.14: Amplification factor versus gap size. λ = 1550 nm, Total Metal Length = 0.5λ = 775 nm, Thickness = 40 nm, Flare angle = 40°, Metal = Gold.

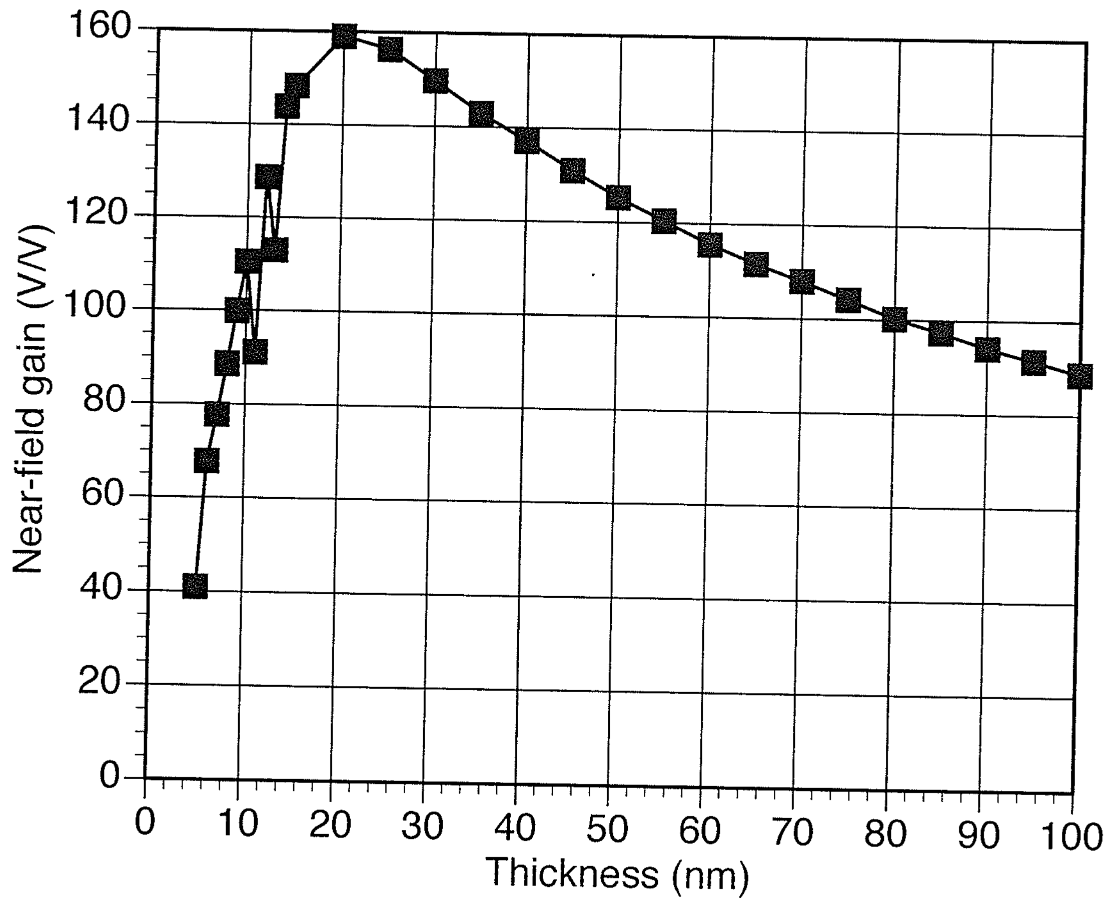


Figure 2.15: Amplification factor versus metal thickness. λ = 1550 nm, Total Metal Length = $0.5\lambda_d$ = 775 nm, Spacing = 10 nm, Flare angle = 40°, Metal = Gold.

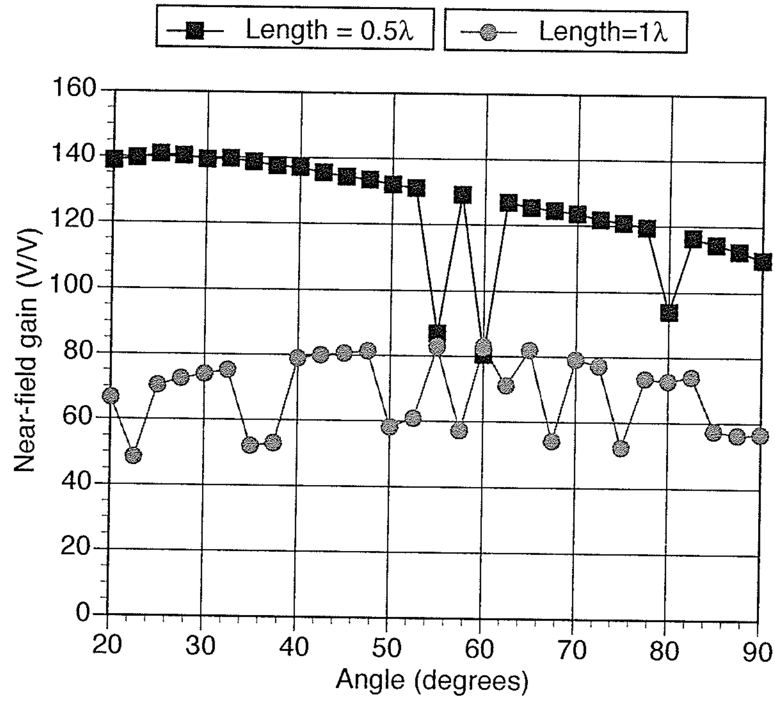


Figure 2.16: Amplification factor versus flare angle. λ = 1550 nm, Total Metal Length = $0.5\lambda_d$ = 775 nm, Spacing = 10 nm, Thickness = 40 nm, Metal = Gold.

nm of the metal surface. Beyond this point the available signal has decreased significantly.

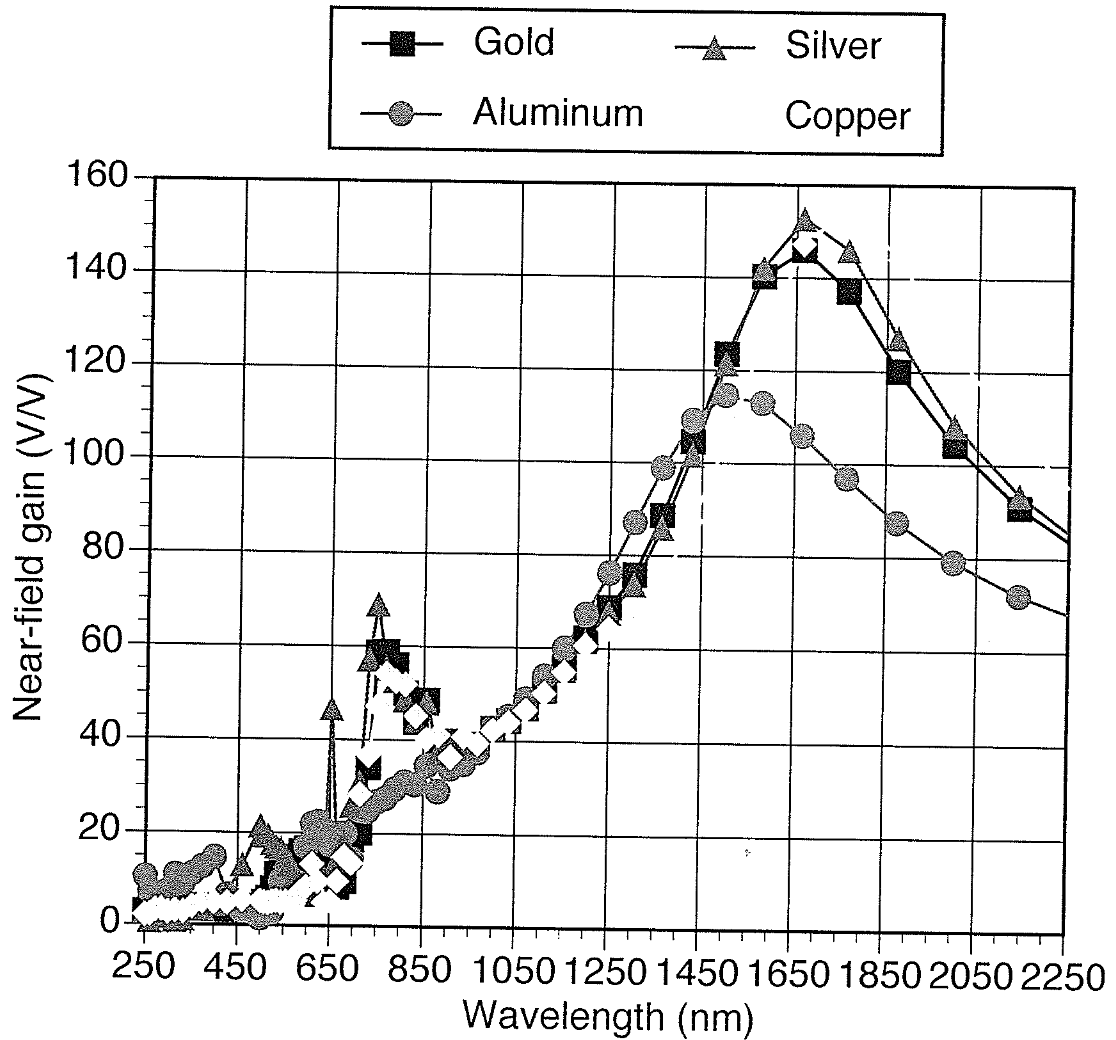


Figure 2.17: Amplification factor versus excitation wavelength for different metals. Total Metal Length = 0.5 λ_d = 775 nm, Spacing = 10 nm, Flare angle = 40°, Thickness = 40 nm.

Chapter 3

Antenna Layout and Illumination Apparatus

3.1 Introduction

The simulation results of Chapter 2 have given us the desired shape of our antennas. In this chapter we will describe the specifics of the antenna structure, the optical patterns used to align a laser to the antenna, and the optical apparatus use to couple the light from the laser to the antenna within a scanning probe microscope (SPM).

3.2 Illumination and Device Layout

One of our key considerations is how to optimally illuminate the samples with an incident beam of light given the geometric constraints of our analysis techniques. Our plan was to analyze the probes using techniques in the SPM family. Since these techniques require a probe to be brought into close proximity to the nanoantennas, a structure that allows illumination from the bottom of the sample is necessary. The small size of the antennas, coupled with the fact that the focused laser beam is also small and, if we are operating at 1550 nm, outside the visible, alignment of this beam to any fabricated antenna structure is a significant challenge. The alignment problem was solved by using a pattern around the antennas that simplified bringing the antennas into the focal plane and positioning the focal point on to them.

Our first consideration is how we will manipulate the focal point of the optical beam. It was decided to make use of a fiber positioner with a lens at the center, so that the focal point may be precisely manipulated in the xy-plane with little mechanical backlash. The height of the sample can then be adjusted to move the antenna into the plane of focus of the laser spot. To bring the beam to the lens we use a collimator attached to the laser. Since supporting a collimator pointing straight to the lens would make the apparatus too high, we instead bounce the beam off of a corner prism. A sketch of the required apparatus is shown in Figure 3.1.

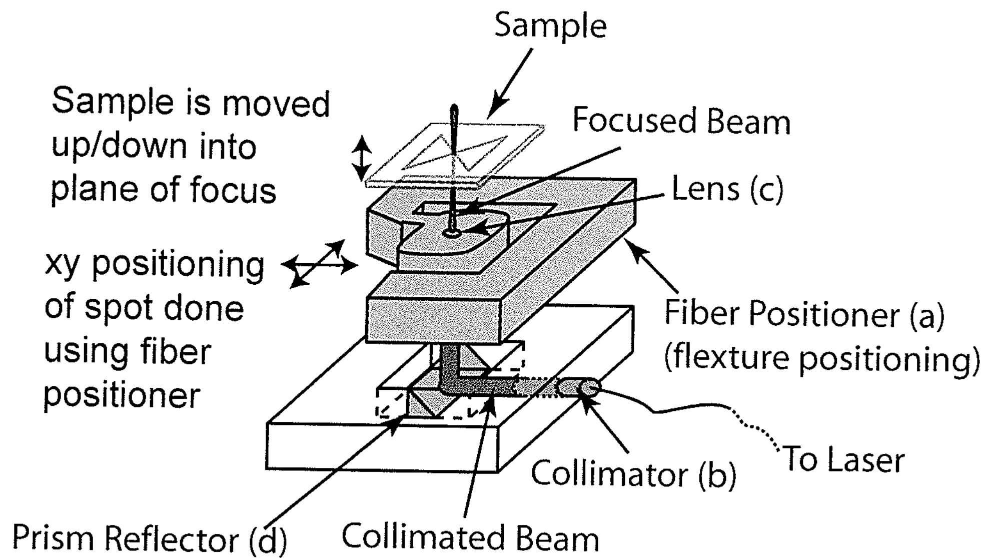


Figure 3.1: Components required for beam alignment.

With this apparatus to control the position of the beam, let us consider the features we have developed for our samples to simplify the process of laser alignment. Since we are illuminating our samples from below, we use the transmitted power, combined with a tapering channel to guide the beam towards the antenna and to bring them into the focal plane. Initially the channel will be wide, making it relatively simple to find a position where the entire beam passes through the substrate to a detector on the opposite side. As the beam approaches the antenna, the channel narrows. By following one of the edges, the tapering features will guide our beam towards the antenna. By monitoring the transmitted power, we can determine when the spot reaches

the apex of the tapered channel. At this point the channel width is the same size as the spot size of the focused beam. Here the channel ends, but shortly after this we place a window with the antennas at the center. A sketch of the tapering channel and its use in aligning the beam to the antenna(s) at the center of the window is shown in Figure 3.2. This sketch does not illustrate the vertical movement of the sample to bring it into the focal plane of the lens. A more quantitative explanation of the laser focusing process is given in Section 3.4.

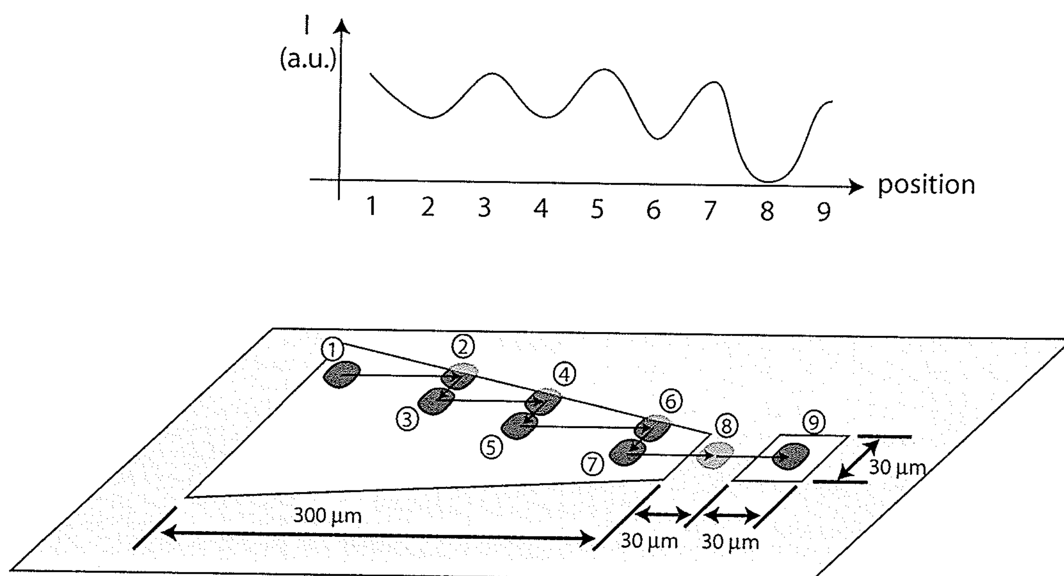


Figure 3.2: Illustration of the tapering channel and its use for beam alignment.

3.3 Illumination Apparatus

For the fiber positioner ((a) in Figure 3.1) we have used a *xy* positioner from New Focus, Inc., part number 9051(M). This device has 3 mm of travel in the *x* and *y* directions, and moves 4.4 μm per 5° turn. The fiber collimator (b) was a Thor Labs 50-1550-FC grin fiber collimator. The specifications for this part, taken from the company website[1], are shown in Table 3.1. The final commercial component in the apparatus is the lens (c) to focus the laser beam. For

[1] http://www.thorlabs.com

this we have used a moulded glass aspheric lens from Thor Labs, model number A390TM-C. The specifications for this lens, taken again from the company website, are given in Table 3.2. Although this lens comes with a holder, in order to fit properly inside the fiber positions another holder was machined. To house the collimator and the corner reflector, and to holder the fiber positioner in place, an aluminum base was milled. The individial components are shown in Figure 3.3, an apparatus to hold these components and a detector to measure the transmitted power are shown in Figure 3.4.

Table 3.1: Specifications for Thor Labs 50-1550-FC grin fiber collimator.

Central Wavelength	1550 nm
Optimized Working Distance	20 mm
Spectral Bandwidth	60 nm
Insertion Loss	Min: 0.15 dB Max: 0.2 dB
Return Loss	65 dB
Acceptance Angle	0.15°
Beam Divergence	0.25°
Beam Diameter	0.5 mm FWHW Max
Optical Power	300 mW Max

Table 3.2: Specification for Thor Labs A390-C molded glass aspheric lens.

Wavelenth Range	1050-1550 nm
Numerical Aperture	0.53
Effective Focal Length	4.6 mm

Having chosen the lens, we can now estimate the size of the focused spot, as well as the depth of focus. Having these numbers will tell use how large to make the window that contains the antennas, as well as give us an estimate as to how difficult it will be to place our substrate in the focal plane of the lensed beam. The spot size (diameter) of a lensed beam is[38]:

$$d = \frac{2.44\,\lambda\,f}{D}, \tag{3.1}$$

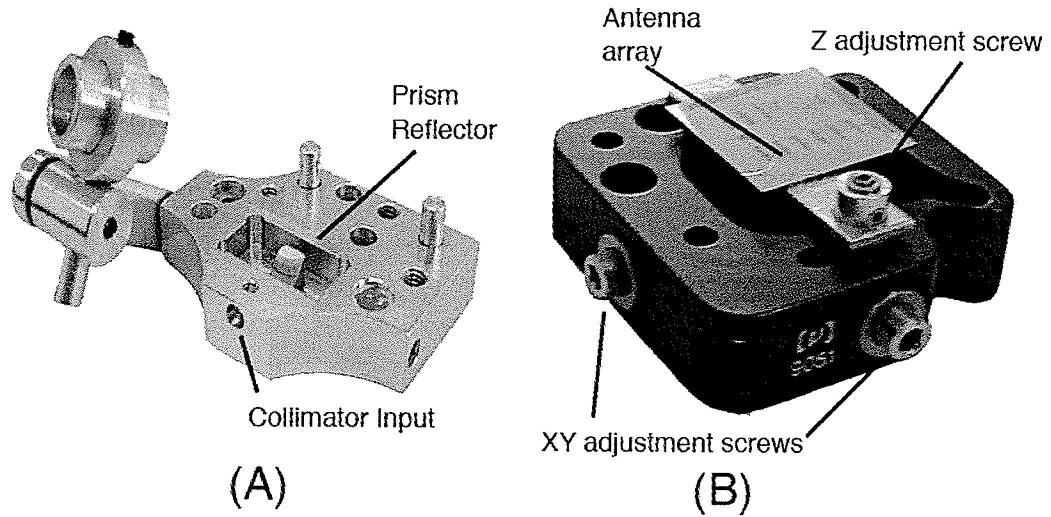


Figure 3.3: Individual apparatus components. (a) Aluminum base with scattered-field detector positioner (b) New Focus 9051(M) Fiber positioner, with sample attached

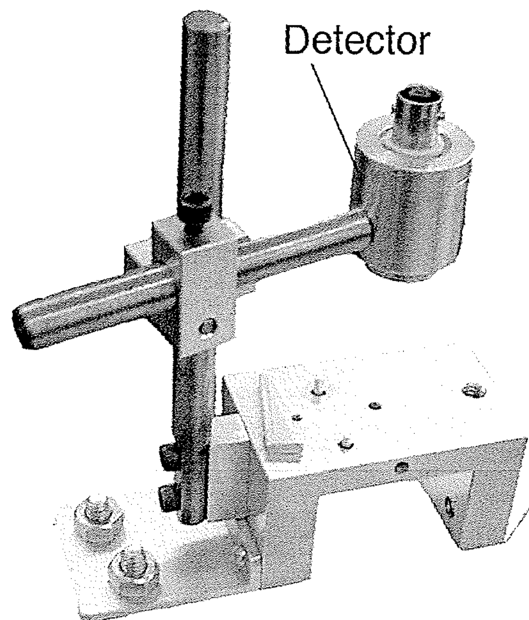


Figure 3.4: Alignment apparatus.

where f is the focal length of the lens and D is the beam diameter before the lens. Using the values in Table 3.1 and Table 3.2, we find the spot size to be:

$$d = \frac{2.44 \times (1550 \times 10^{-9}\ 4.6)}{0.5 \times 10^{-3}} = 34.8\,\mu\text{m}, \tag{3.2}$$

From the spot size we can determine the depth of focus. This is[38]):

$$dof = \frac{\pi\, d^2}{2\lambda}. \tag{3.3}$$

For our lens we have

$$dof = \frac{\pi\,(3.48 \times 10^{-5})^2}{2 \times 1550 \times 10^{-9}} = 1.23\,\text{mm}. \tag{3.4}$$

From the values calculated in 3.2 and 3.4, we determine that the window should be roughly 30 μm × 30 μm (so that we can assume a mostly constant field inside) and that is should be relatively straightforward to ensure the sample lies in the focal plane. When we align the laser in the next section we see that our calculations are consistent with our measured results. It should be noted that the spot size is 24.4× the wavelength. If we could better focus the light, we would obtain a higher intensity from the antennas, but alignment would be more difficult. Additionally, achieving a better focus would require larger optics, which we do not have space for in our apparatus.

3.4 Laser alignment

The process of aligning the laser spot to the antennas is difficult and requires practice. The first challenge, which is also the most difficult, is to attach the sample in the correct position. A visual representation of this process is given in Figure 3.5.

The process begins by adjusting the positioner so the lens is at the center of the xy-range (and the center of the beam opening) (step 1 in Figure 3.5). From here the detector should be positioned over top of the lens, and the user should ensure that sufficient xy travel (at least a full two turns in any direction) is available. The lens is then moved to one extreme, depending on the orientation of the tapering channel (step 2). The extreme is the point where the transmitted

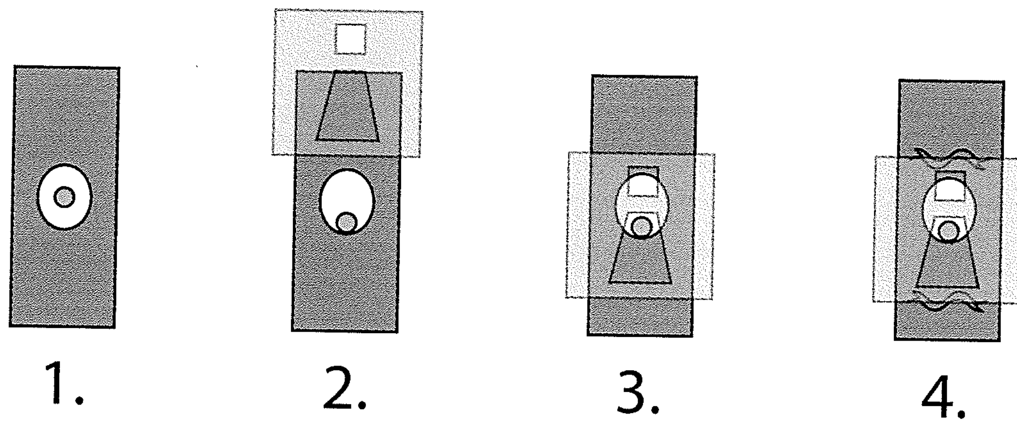


Figure 3.5: Procedure for gluing samples to ensure laser alignment.

power is still maximal, but would drop if moved more than a full turn. The sample should then be placed on the beam over top of the lens and positioned in such a way that the beam is as far down the channel as is possible (step 3). This step requires the user be very precise, but will ensure that there is sufficient travel to position the spot inside the window. Sample glue (Loctite 454) is carefully and sparingly applied to affix the substrate to the beam (step 4).

Once the sample has been properly glued, the focusing procedure shown in Figure 3.2 can be performed. The transmitted power will drop to near zero in the region between the end of the taper and the window (step 8 in Figure 3.2). Once the beam is in the window, the power transmitted should be very close to what was transmitted in the channel (within 10%). At the center of the window it should only require 1/24th to 1/16th of a turn in any direction to reduce the transmitted power to near zero. Typical values seen during the focusing process are give in Table 3.3. It should be noted that the collimator is very sensitive to how it is attached to the laser, and several tries may be necessary to obtain maximum output. After the spot is inside the window, the z-postion can be easily fine-tuned, but until the z-postion is near the focal plane it will be difficult to observe a significant drop in transmitted power when the beam is between the channel and the window. The adjustment in the z-direction should thus be done iteratively once there is evidence the end of the channel is near. The large depth of focus of our system makes this process

Table 3.3: Typical detector measurements for 100 μW input laser power.

Collimator→Detector	.042 μA
Collimator→Channel→Detector	.036 μA
Collimator→Between channel and window→Detector	.001 μA
Collimator→Window→Detector	.032 μA

easy. Note that the detector may need to be slightly repositioned during the alignment process to get a more accurate measure of the transmitted power.

3.5 Scattered Field Detection Apparatus

For the experiment in Section 5.3 we perform measurements using a system like that of apertureless NSOM. For the experiment we make use of a lens and photodiode to collect as much of the scattered field as possible. The lens used is Kodak designed moulded glass aspheric lens, part number A397TM-C. The specifications for this lens are given in Table 3.4. The detector is an

Table 3.4: Specification for Thor Labs A397TM-C molded glass aspheric lens.

Wavelenth Range	1050-1550 nm
Numerical Aperture	0.30
Effective Focal Length	11 mm

InGaAs photodiode from Thor Labs, part number FGA21. Figure 3.6 shows the spectral response between 800 nm and 1800 nm. The electrical characteristics of this diode are given in Table 3.5. The photodiode outputs a current, so in order to provide a signal to the lock-in amplifier, a 2.2 kΩ resistor was attached in parallel prior to the input of the lock-in. Both components were attached together using a milled sleeve 11 mm long (the focal length of the lens). A positioning apparatus was attached to the base, which allows the detector to be pointed towards the antennas. The lens should be approximately 11 mm away from the antennas, which would allow for maximal collection of the scattered light, since this would focus a point near the antenna to a point on the detector, as shown in Figure 3.7. The detector attached to the illumination apparatus is

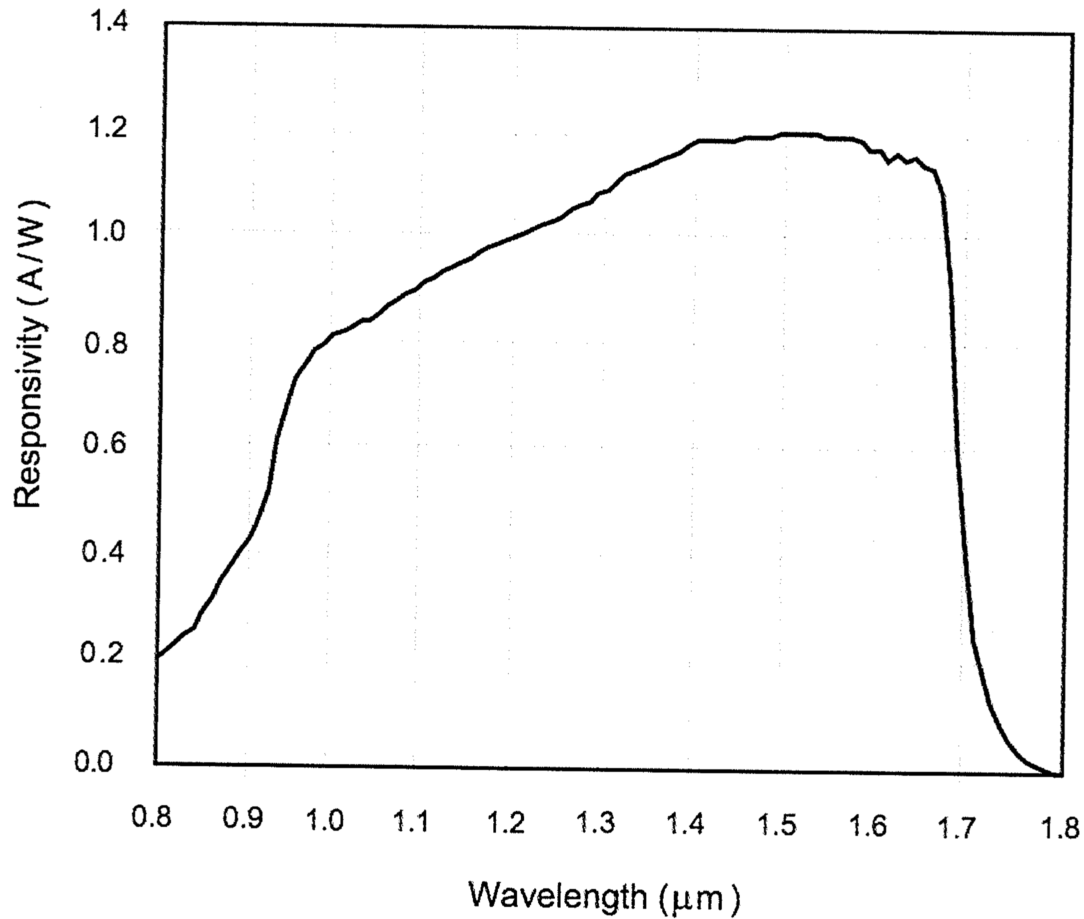


Figure 3.6: Spectral Response for FGA21 photodiode.

Table 3.5: Specifications for Thor Labs FGA21 InGaAs-Photodiode.

Spectral Response:	800 - 1800 nm
Active Area:	Ø2mm (3.1 mm^2)
Rise Time (R_L = 50Ω):	66 ns (0V bias)
Fall Time (R_L = 50Ω):	66 ns (0V bias)
Dark Current:	200 nA max. (1V)

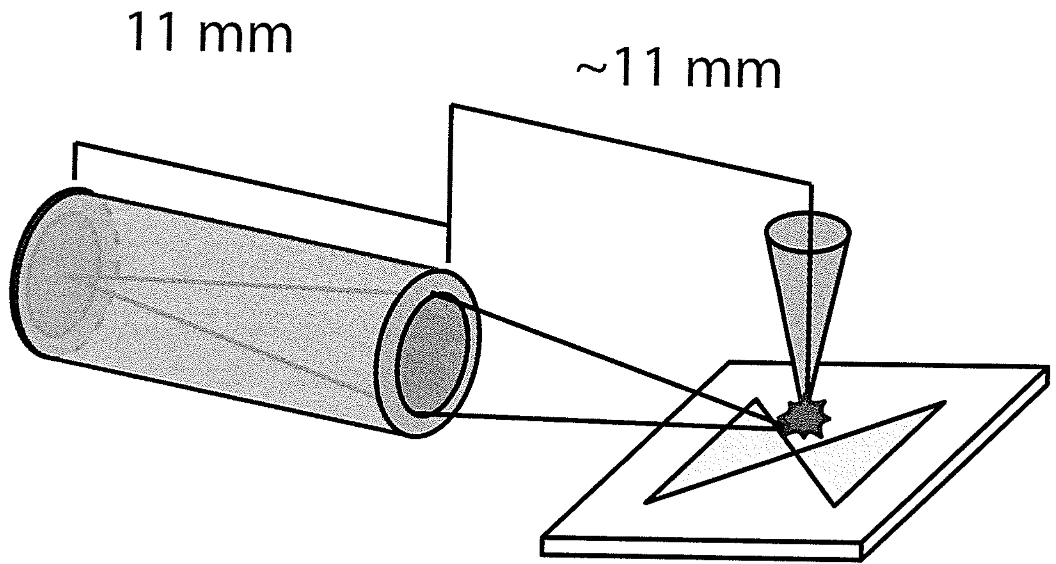


Figure 3.7: Focusing scattered field detector.

shown in Figure 3.8.

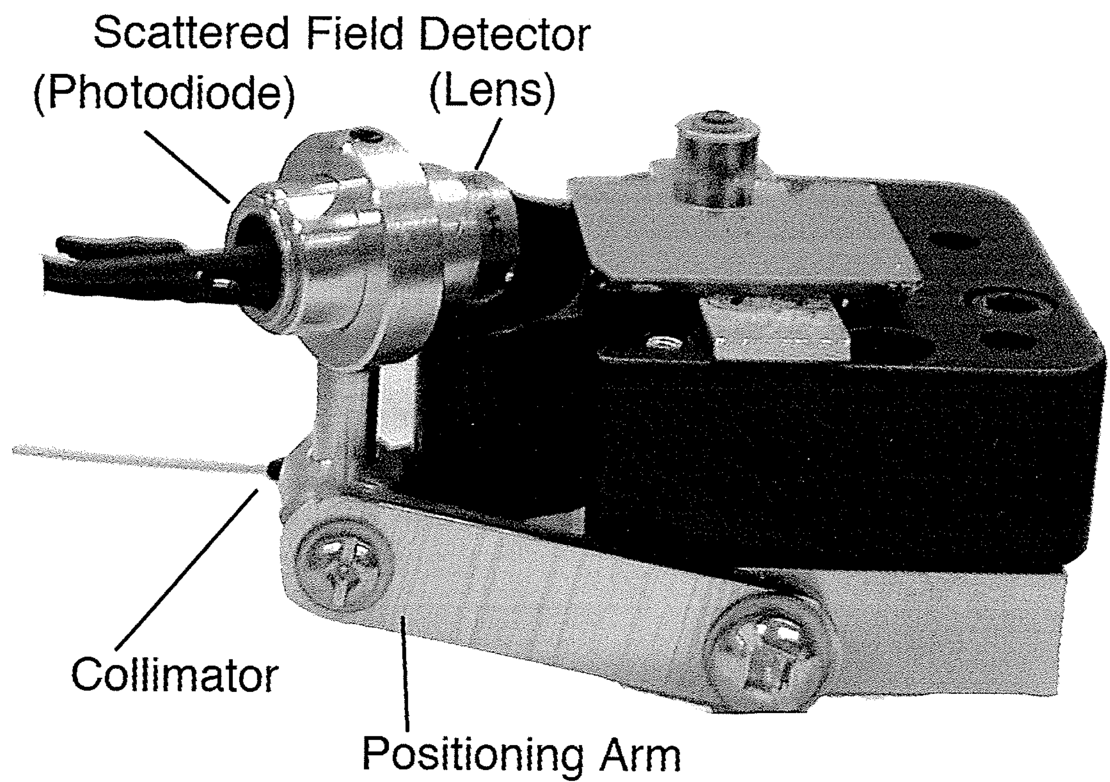


Figure 3.8: Illumination apparatus combined with scattered field detector.

Chapter 4

Fabrication

4.1 Introduction

The simulation results of Chapter 2 have given us a guideline for the antenna geometry. This chapter deals with the challenge of device fabrication. Our simulation results show that for the antennas to act as effective transducers for 1550 nm wavelength light the separation between the two probe halves should ideally be below 20 nm. Figure 4.1 shows the desired geometry. These requirements severely restrict the potential fabrication methods. In particular, these features are well below the limits of conventional optical lithography. In this section we will examine potential fabrication materials and techniques, and present the results obtained using these techniques for creating the antennas.

4.2 Materials

Because of the high frequencies and small dimensions involved in these devices, we must be careful in our selection of the metal for the probe body as well as the substrate on which the probe rests. We would like our metal layer to satisfy the following criteria:

- It must posses suitable optical properties.
- We must be able to pattern it using the chosen fabrication method.
- We would like the chosen metal to be relatively inert.

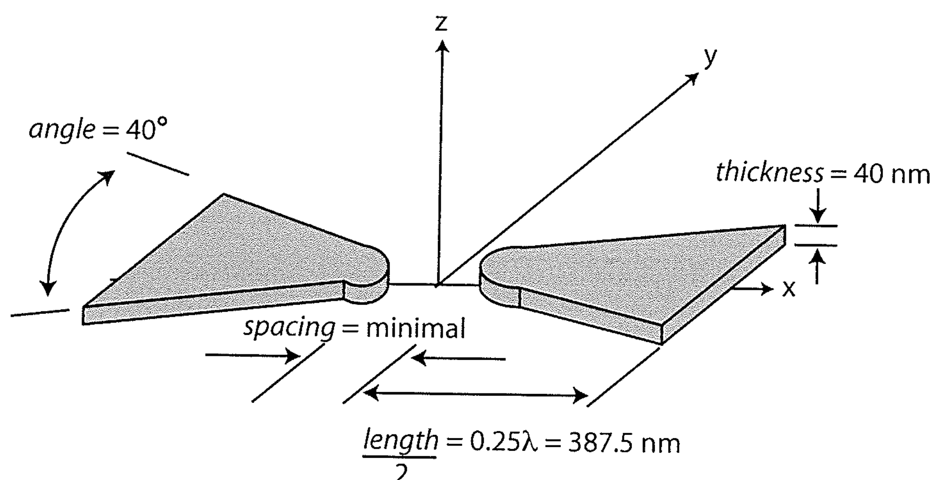


Figure 4.1: Desired antenna geometry.

We reviewed optical properties in Section 2.2, and determined that suitable metals included gold, silver, copper and aluminum. Of these, gold was chosen for our final fabrication due to its inertness, ease of use in fabrication, and favourable optical properties.

The choice of substrate material must consider the illumination geometry of the experiment. As shown in Section 3.2, we will be illuminating our sample from below. A suitable substrate should thus be transparent to 1550 nm radiation. In order to be able to use the same procedure for illumination with visible light, we have chosen to perform our fabrication on glass slides.

Indium Tin Oxide (ITO) coated glass slides 0.5 mm thick, part number CB-90IN, were purchased from Delta Technologies[1]. The glass layer is made of display grade corning 1737, and the nominal coating thickness is between 150 and 300 $\AA$, which gives then a sheet resistance R_S between $70 - 100\,\Omega/\Box$ and a nominal transmittance above 85%. The ITO coating reduces charge buildup during ion milling and electron beam lithography, which should increase the resolution of the small features since built up charges would scatter the incoming ions.

[1]www.delta-technologies.com

4.3 Deposition

Prior to metal deposition all slides were cleaned using a three step process which consisted of washings in acetone, followed by isopropyl alcohol, followed by methanol. The slides were dried using a N_2 gun following each step. Following cleaning, the slides were placing inside a MRC 8667 sputter system from Semicore[2] for metal deposition. To promote adhesion between the ITO layer and the gold layer, a thin layer of approximately 5 nm of Cr was first sputtered. This was done with a DC sputter at 200 W. The chrome was sputtered for 5 seconds after the turn-on time. Following this, approximately 80 nm of Au was deposited at the same power as the Cr sputter, sputtering for 48 seconds. While it was estimated that this would deposit roughly 40 nm of Au, when measured the total thickness of the metal was close to 90 nm. Although this is not ideal, it was determined from the simulation results that this would not significantly affect the operation of the antennas. Following the Au deposition it was noted that its inertness makes it difficult for photoresist to adhere. This can make it difficult to achieve high-resolution lithographic feature. To fix this problem, a final 5 nm thick layer of Cr was sputtered on top of the Au (using the same process as the previous Cr layer) to act as an adhesion promoter.

4.4 Device Layout Consideration

For our proposed experiment, we need another requirement from potential fabrication techniques. Since our experiment requires us to focus a laser on the antenna, we have designed the device in such a way that laser alignment is simplified. Additionally, since laser alignment is a time-consuming process and the laser cannot be moved once the apparatus is inside the vacuum chamber, we fabricated multiple antennas close to one another. As discussed in Section 3.2, our spot size is roughly 30 μm, so multiple antennas can be illuminated at one time, if they are fabricated close enough to one another. Given the list of requirement for the fabrication of nanoantennas, two techniques were explored for device fabrication: focused ion beam milling and electron beam

[2] http://www.semicore.com

lithography. The advantages, disadvantages, implementation and results of both techniques are discussed below.

4.5 Focused Ion Beam (FIB) Milling

4.5.1 FIB Technology

The focused ion beam (FIB) is a tool that is designed for milling very precise designs. They are most commonly employed to trim magnetic read/write heads and to cross-section materials for inspection and failure analysis in the microelectronics industry[39].

Apparatus

The main component inside any FIB system is the ion column which emits, focuses and directs a beam of energetic ions towards the sample to be imaged and/or milled. Figure 4.2 shows the components inside a FIB column. The apparatus is very similar to what is used inside a scanning electron microscope, with the main difference being that, in the place of an electron beam, a beam of gallium ions is used (Ga^{+}). A liquid-metal ion source is used to generate the beam. A liquid gallium cone is formed on the tip of a tungsten needle, and a strong electric field (typical extraction voltage is 7000 V, applied through the extractor creates the ion beam. Under typical operating conditions the extraction current is roughly 2 μA. The suppressor pushes back most of the ions through the use a strong reverse bias field. Only the ion traveling in a straight line near the center of the suppressor can pass.

The initial beam refinement is done with the spray aperture, and the first electrostatic lens then condenses the beam. The upper octopole is used to adjust the beam stigmatism, and the variable aperture is then used to provide a range of beam intensities from very fine (for high resolution imaging and precision milling) to very broad (for rapid milling). Stigmation occurs due to a lens having two line foci instead of a single focal point. This causes the image to be appear "streaked" in a certain direction, and the upper octopole is used to compensate. The variable aperture allows for a large beam-current variation, between roughly 1 pA and 10 nA. A typical system allows a choice of

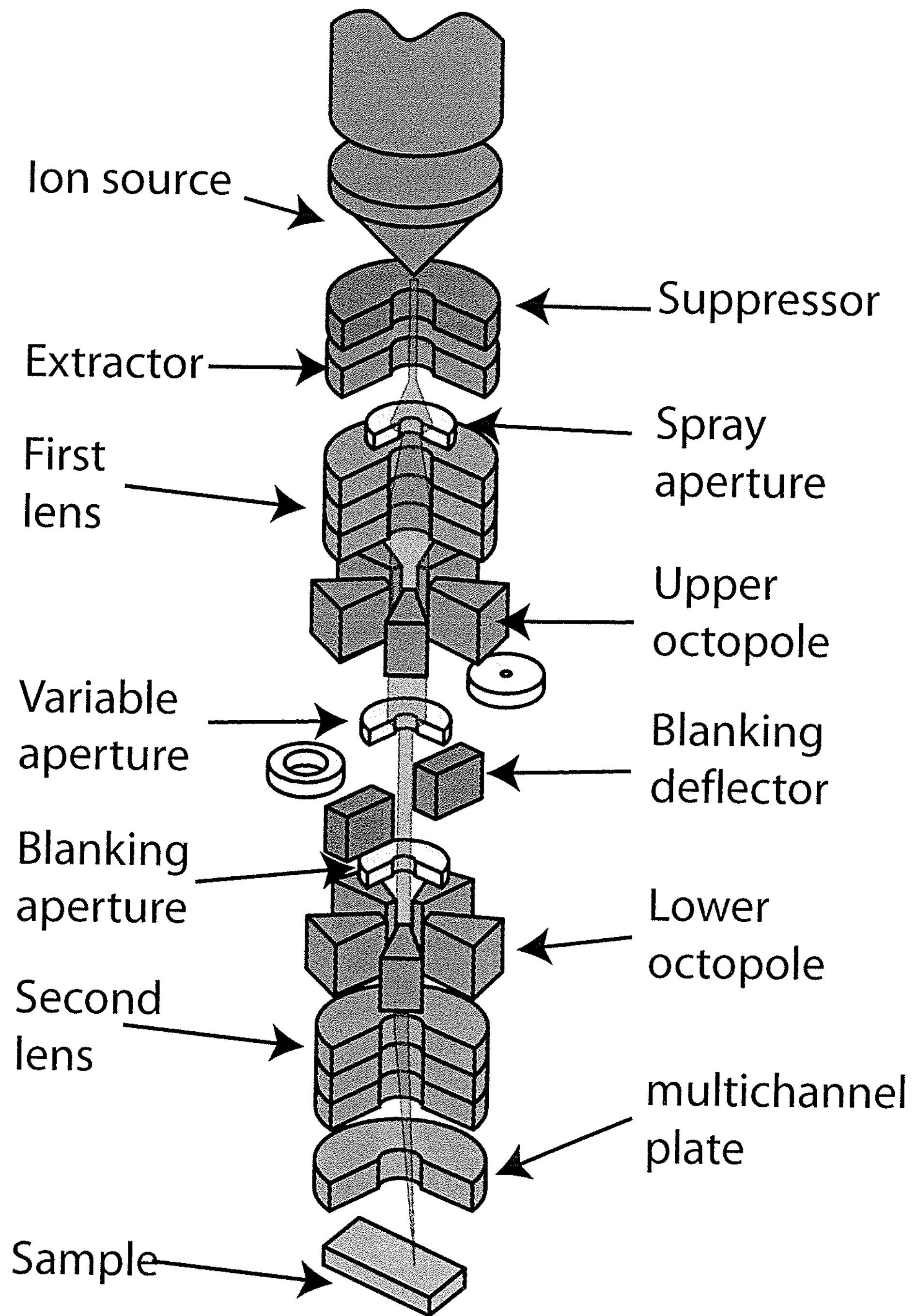


Figure 4.2: Components of ion column. Reproduction of Figure 2 in [39]

seven values for the beam current inside this range.

The next stage in the ion column is the two part blanking stage. This allows the beam to be rapidly shut off by deflecting the beam off of the target. If the beam continues, then lower octopole and the second second electrostatic lens focus the beam to a very small spot. If we are operating in imaging mode the multichannel plate (MCP) collects the secondary particles.

The ion column must maintain a low pressure (1×10^{-7} mbar) for proper operation. The ion beam is then directed into a vacuum pumped work chamber, which must also be kept at a low pressure (low 10^{-7} mbar range) to ensure a high mean free path thus maintaining the tightly focused beam. Most FIB systems use a load-locking system to ensure that changing samples can be done with as little loss of vacuum as possible. The vacuum system inside the work chamber often allows for other gasses to be introduced near the sample. This can be used to improve both the rate and selectivity of the ion milling, as well as to deposit materials (described in the next section).

Imaging, Milling and Deposition

The energetic ions ejected from the ion column can interact with the surface in a variety of ways. The nature of the interaction is dependent on several parameters, most importantly the beam energy and the physical and chemical nature of the incident surface. Figure 4.3 shows the different modes of operation that result from various interactions: imaging, milling, and deposition.

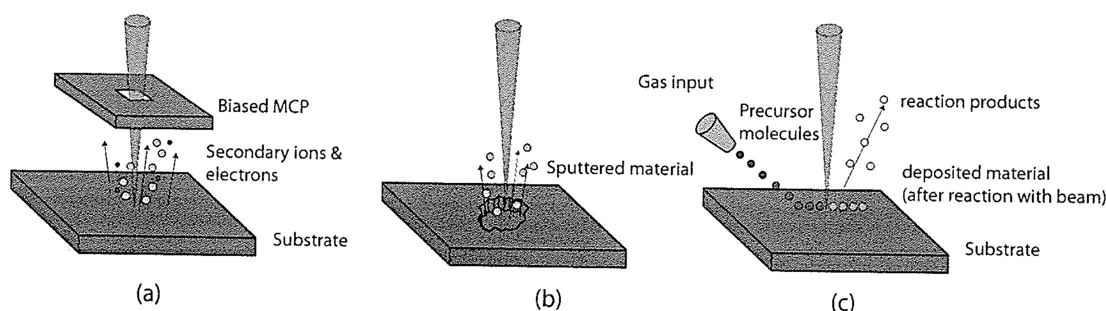


Figure 4.3: Modes of operation of the focused ion beam. (a) Imaging, (b) Milling, and (c) Deposition

Figure 4.3a shows the FIB operating in imaging mode. The column is set to emit a tightly focused, low current beam (roughly 1-30 pA) which is raster

scanned over the surface. The excitation from the beam causes the emission of secondary particles, which are collected by the biased MCP. A positive bias on the MCP will collect secondary electrons, while a negative bias collects secondary ions. With an attached mass spectrometer, secondary ion mass spectroscopy can be performed. The ultimate resolution of the imaging process is determined by the beam spot size, and modern units have beam diameters of a few nm[39].

There are some difficulties associated with imaging using an ion beam. Some implantation of Ga^+ ions is inevitable, and when combined with the secondary electrons leaving the surface (regardless of if they are being detected) can cause buildup of positive surface charge. If the imaging is being performed using secondary ions, then an electron source can be used to negate the charge buildup. This allows for imaging on non-conductive samples. Another issue with the imaging process is that milling of the sample is inevitable. Many FIB systems are also equipped with an electron column, as SEM imaging is less damaging to the surface. The SEM column will be at an angle however, so care must be taken if the SEM images are used to plan milling locations.

Figure 4.3b shows the FIB operating in milling mode. A high current beam is directed over the areas to be milled. This physically removes material from the substrate by sputtering. While the process is relatively straightforward, it is very difficult to calculate the rate of material removal *a priori*. While typical removal rates are known, in practice the removal rate is dependent on the pattern being etched. This is mostly due to re-deposition of some of the material sputtered from one region into another etch location. Both the selectivity and the etch rate can be improved through the introduction of a particular etching gas in to the work chamber.

Figure 4.3c shows the final mode of operation of a FIB, deposition. This is the least common mode of operation. The process is a type of chemical vapour deposition. Gas is adsorbed on to the surface and the incident ion beam then decomposes the gas. This removes the volatile reaction products and leaves behind a thin layer of material. Commercial systems can allow for deposition of platinum (Pt), tungsten (W), SiO_2 and oxygen (O_2)[39]. This work did not make use of this mode.

4.5.2 FIB Fabrication procedure

While the FIB beam is well suited to creating very small features, removal of large amounts of material is better done by other means. We have chosen to use traditional optical lithography to create the tapering channels using to guide the beam towards the bowties, as well as to create an isolated metal island inside a window. This island is then etched using the FIB to create the nanoantennas. The entire fabrication processes of deposition, optical lithography and focused ion beam milling is shown in Figure 4.4

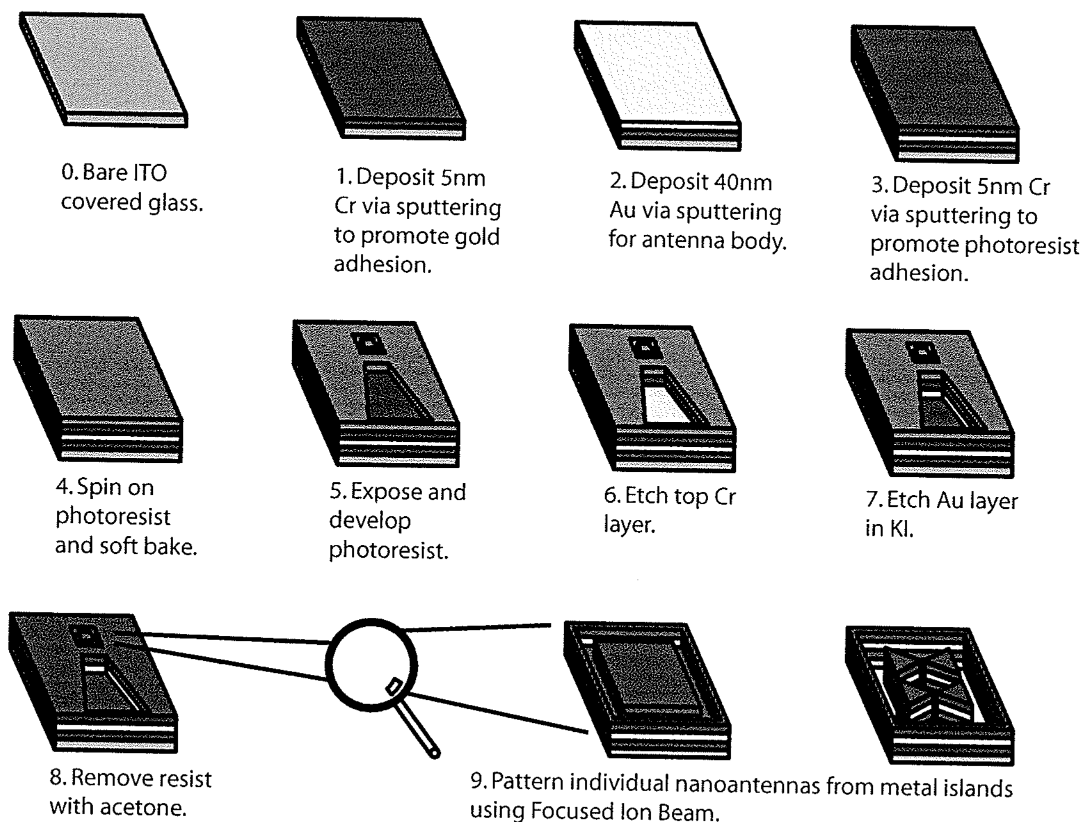


Figure 4.4: Fabrication with focused ion beam.

Optical Lithography

Steps 1-3 in Figure 4.4 are the metal deposition steps covered in Section 4.3. The stages of optical lithography are shown in steps 4-8 inclusive. Beginning with the ITO slides covered with the Cr-Au-Cr film, HPR 504 photoresist was spun on to the slides at 3000 RPM for 60 seconds. This creates a 1.5 μm thick layer of photoresist. Immediately after this the slides were placed in a softbake at 110°C for 60 seconds. The slides were then transferred to a mask aligner and exposed for 8 seconds. The mask for this step was designed in the software program Virtuoso by Cadence[3] and fabricated at the University of Alberta. Figure 4.5 shows the mask layout. Following exposure the photoresist was then developed in 352 developer for 12 seconds and hardbaked for 20 minutes at 110°C. Following developing the top Cr layer was etched for a few seconds using a commercial Cr etchant purchased from Arch Chemicals. Then samples were washed and dried, and the Au layer was removed using a Potassium Iodide (KI) etch. To prepare the etchant 4 g of solid KI and 1 g of solid I_2 were measured using a digital balance and placed in a glass container. 40 ml DI water was added and the solids were dissolved by agitation overnight. The resulting solution was used to etch the gold layer by placing the sample in the solution for approximately 8 seconds, then removing, washing and drying the samples. The final Cr layer was not etched because of the possibility that over-etching could detach the islands from the substrate, and because the FIB would have little difficulty in removing the extra metal. The remaining photoresist was removed using acetone. Figures 4.6 and 4.7 show optical micrographs of the resulting structures.

Post-Processing using a FIB

Following the completion of the optical lithography steps and verification of the approximate shape of the metal islands, the samples were brought to the FEI[4] Strata DB235 focused ion beam system in the *laboratoire de microfabrication* at the *École Polytechnique de Montreal*. The unit contains both an ion column and an electron column, which allows for both FIB and SEM imaging and milling. The sample is first inserted into the system using a load-lock,

[3]www.cadence.com
[4]http://www.feicompany.com

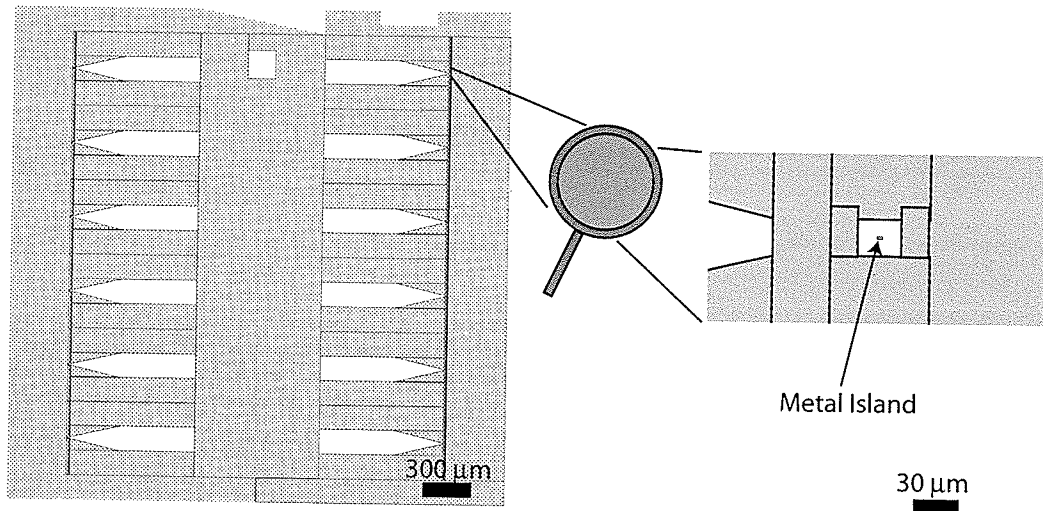


Figure 4.5: Mask layout. The windows on the left side of every quadrant are 60 µm square, while the windows in the right are 30 µm square. The size of the islands vary in size. From the top line to the bottom, the sizes are: 1 μm^2,2 µm × 1 µm, 2 μm^2, 3 µm×2 µm, 3 μm^2, and 4 µm × 3 µm (on both the left and right sides).

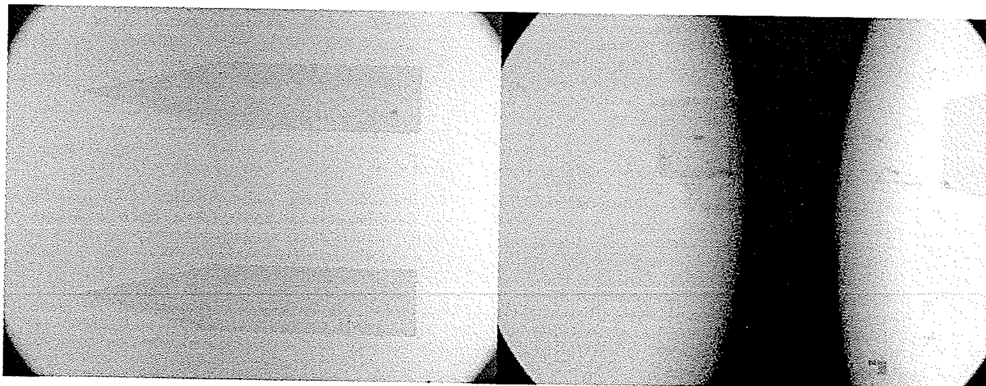


Figure 4.6: Optical micrograph of right side window and tapering channel at 10× and 100× magnification.

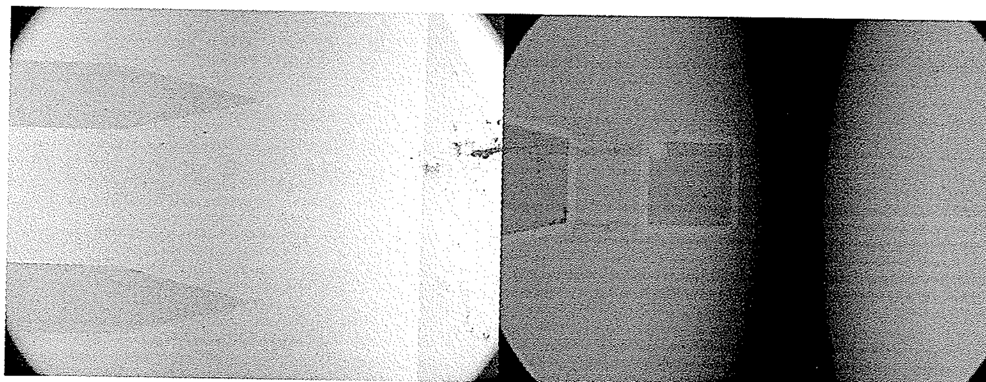


Figure 4.7: Optical micrograph of left side window and tapering channel at 10× and 100× magnification.

and the chamber is pumped down. After initializing the beams, SEM imaging mode is engaged. After the sample is brought in to focus, the height must be adjusted so that both beams point at the same location, as shown in Figure 4.8. This is known as the *eucentric height*, and should be checked again prior to beginning the milling process. Once the desired feature is brought in to focus, a single image is taken using the ion column, the pattern is drawn on the screen, and the pattern is etched using a 10 pA beam intensity. Because material is re-deposited during the milling process, the order in which material is removed is important. Figure 4.9 shows the pattern that was used to create the antennas. The final three steps were the most crucial. Failure to place these three steps at the end, in this order (though steps 6 and 7 can be reversed), will likely result in material being re-deposited in the gap region.

There are a few parameters that, in theory, should give control over the line thickness. The ion beam software allows the user to set both the dwell time, which is the amount of time the beam stays focus on a single spot, and the overlap percentage, which is the amount each point overlaps with the previously drawn one. These parameters are important because of the nature of the patterning software. The drawing tools are all vector based, but in order to translate the input pattern in to commands for control of the ion beam, the image is converted into points. While it would seem that by adjusting these two parameters one could control some aspect of the line thickness, in practice it was found to have no effect for this application. One unexpected parameter that did affect the milling process significantly was the magnification level

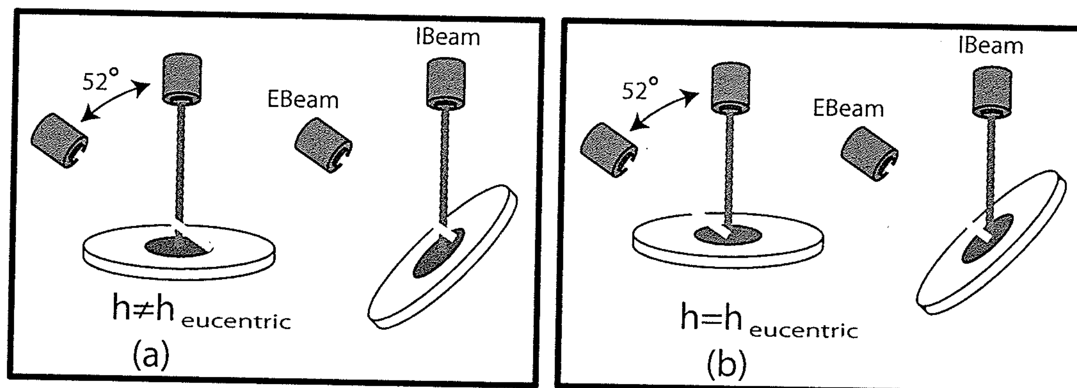


Figure 4.8: Illustration of beam alignment in a FIB system. Note that at the eucentric height, the electron beam will point at the same location at the two tilt positions.

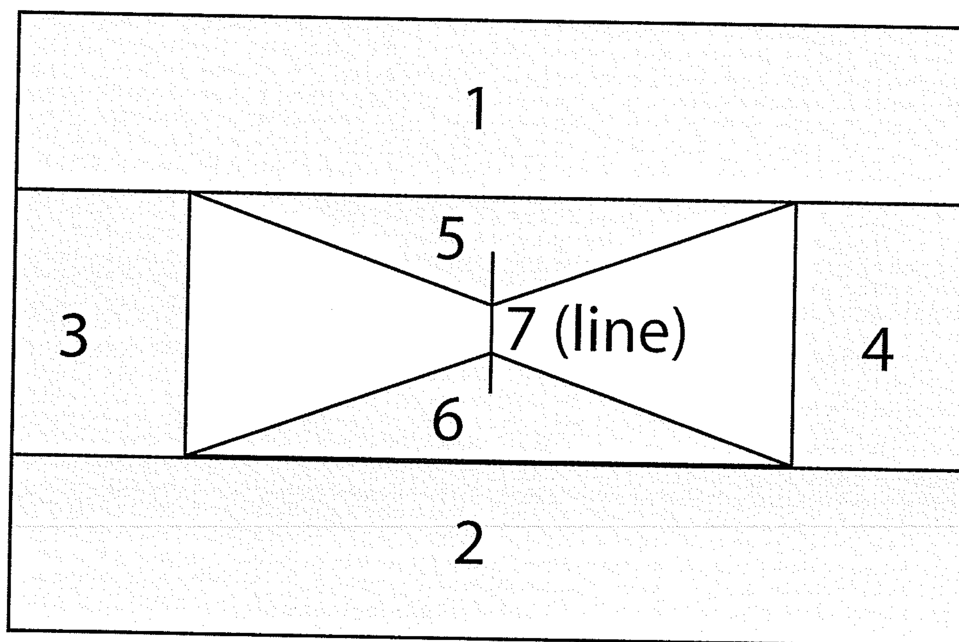


Figure 4.9: Order of Ion beam milling.

of the imaging mode of the ion beam. This is again likely due to the process of converting a vector image in to commands to control the ion beam. The thinnest lines (determine by visual inspection of the SEM image) were obtained at magnification levels of 100000× and 200000×. The highest magnification level possible with the Strata DB235 is 500000×, however this did not result in thinner lines. The reason for this is not know, but it is likely once again related to the conversion between the vector image and the ion-beam control signal.

The removal of material during milling could easily be verified by measurement of the backscattered secondary ions, which decreases drastically once the metal layer has been removed. While prior to milling it was thought that the metal layer was only 50 nm, inspection following FIB milling showed the metal thickness to be closer to 90 nm. Despite this error, measurement of the backscattering indicated that all the metal was removed.

4.5.3 FIB Fabrication results

The results of the FIB fabrication for a typical slide window are show in Figure 4.10 a-e. The outer circle in 4.10a is thought to be the result of over-etching the gold layer. Over top of the of the gold island sits a very thin layer of Cr. The greater thickness of the Au layer made it more difficult to etch without undercut, but the chrome layer remained closer to mask size of the lithographic features. In 4.10b&c we see the variation between multiple antennas milled with identical parameters. There are significant differences in the size of the gaps, likely due to thermal drift and charging effect in the glass layer. Additionally, the effects of thermal drift can be seen in the poor alignment of the antennas relative to one another. The pattern for all of the antennas was input at one, with an uniformly spaced 4×2 grid, but, due to the serial nature of the FIB, drift during the milling process caused an uneven alignment.

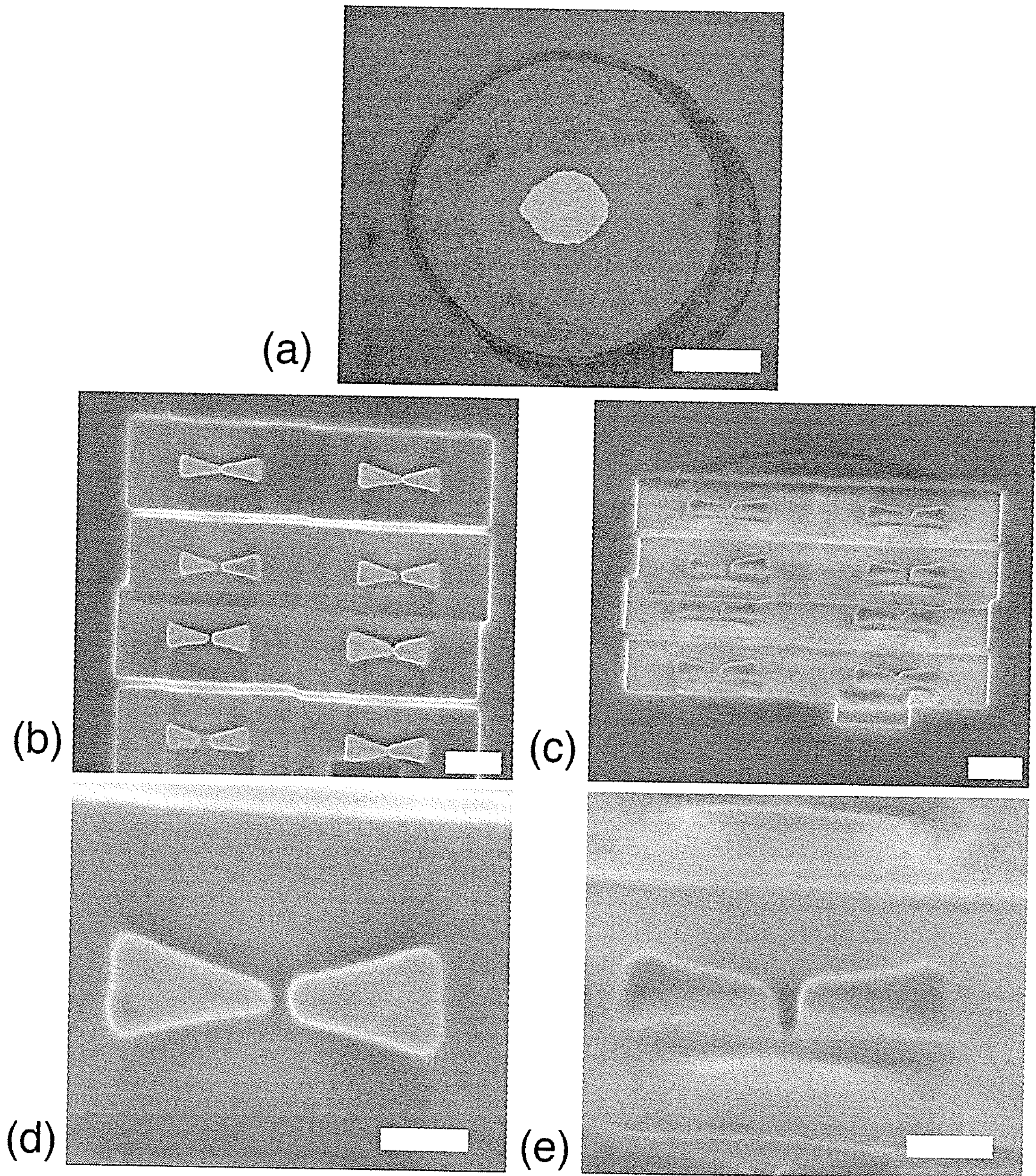


Figure 4.10: Results of focused ion beam fabrication. SEM images showing (a) Metal island prior to fabrication [Scale bar = 1 μm] (b)(c) 8 antennas fabricated from island [Scale bar = 500 nm] (d)(e) Close-up of antenna 2nd row right [Scale bar = 200 nm].

4.6 Electron Beam Lithography (EBL)

4.6.1 EBL Technology

Electron beam lithography is a high resolution technique that bears similarities to both traditional optical lithography and focused ion beam milling. EBL is a direct-write technique, and can be used both for milling patterns directly into semiconductor surfaces, or for exposing patterns on resists to be etched afterwards. It is the latter process which provides most of the benefits to EBL. Because the beam size is incredibly small, the resolution limits are usually a function of the resists used, unlike traditional optical lithography.

Electron Beam Lithography Apparatus

The principle component of an EBL system is the electron emission source. A large variety of different sources are available, using three basic mechanisms for electron emission: photoemission, thermionic emission, and field emission. The electron beam sources used in EBL are most commonly either thermionic emitters or field emitters. [40] discusses the properties of commonly used sources.

Photoemission uses energy from an incident photon to excite electrons in the cathode. The incident kinetic energy can impart enough energy to overcome the work function.

Thermionic emission uses heat to allow electrons to escape. A source that uses this principle contains a negatively charged cathode which in heated to emit electrons. The positively charged anode draws the emitted electrons. Both these components are enclosed in a vacuum, otherwise the electrons would lose their energy traveling through the gas that would be between the two terminals. This also prevents contamination of the cathode via reactions with outside gases.

Field emission, also know as *cold emission*, makes use of a large potential difference between the anode and cathode. This difference reduces the potential barrier seen by electrons in the cathode and allows them to escape through tunneling.

Other than the emitter, the components of an electron column are similar

to those of an ion column. Rather than directly milling a sample however, electron beam lithography is typically used to expose a resist.

4.6.2 EBL Fabrication procedure

The fabrication of the nanoantennas was done by Peng Li at Applied Nanotools[5] in Edmonton, AB using a Raith 150[6]. Table 4.1 shows the specifications for this system. Three different procedures were considered for use in the fabrication of the nanoantennas.

Table 4.1: Raith 150 Specification.

Filament Type	Schottky thermal field emission
Beam size	≤ 2 nm at 20 keV
Beam current range	5 pA - 20 nA
Beam energy	100 eV - 30 keV
Current density	≥ 7.500 A / cm^2.
Minimum line width	≤ 20 nm

In order to fabricate both the large tapering channel with the window at the end, and the small antennas, two sets of steps were required in each technique. One set to create the tapering channel with the window at the end, and the other to create the antennas themselves. Optical lithography could have been used for creation of the channel and window, but it was decided to do all fabrication using the electron beam.

The first method considered is very similar to traditional optical lithography, and is shown in Figure 4.11. In this procedure the metal is deposited and then photoresist would be spun on. The resist is exposed using the electron beam, and the areas where the resist was removed are then be etched. The major predicted difficulty in this process was assumed to be in the etching step. The metal layer is several times thicker than the desired gap size, so wet etching techniques are not feasible since undercut would drastically increase the gap size. An strongly anisotropic etching technique such as RIE or ion milling would be required to etch the metal layer. This approach was deemed to be

[5]www.appliednt.com
[6]www.raith.com

overly complex, since the suitable equipment was not available. It should be noted that the electron beam is quite capable of creating very accurate patterns in resist, as shown in Figure 4.12. Despite the complexity of this technique, it should be capable of producing very well defined antennas, and may be worth exploring. Note that since this technique was not developed further, Figure 4.11 does not show the two-stage process that would be necessary for creating both the alignment features and the antenna.

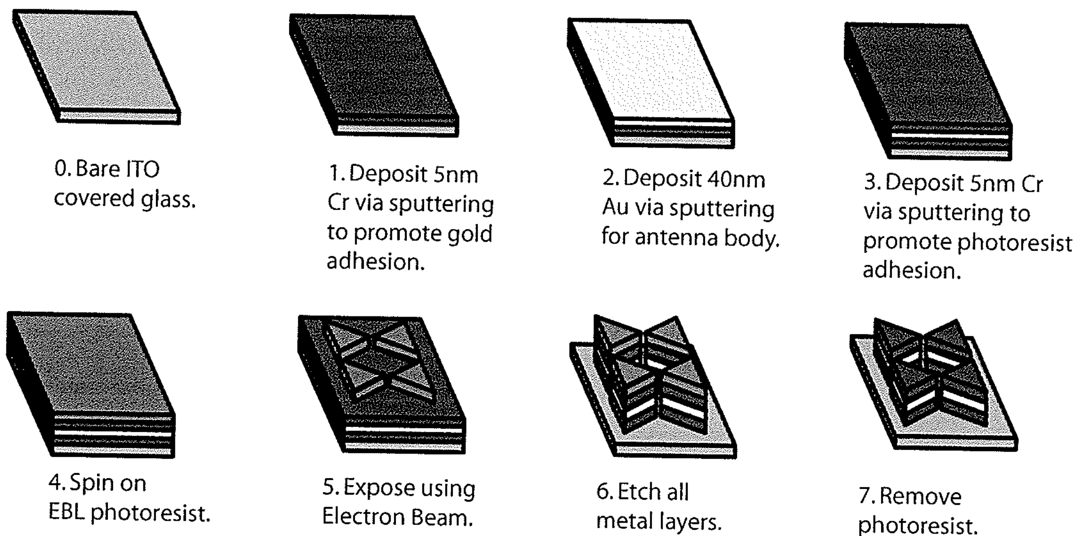


Figure 4.11: First posible EBL fabrication technique.

The second and third procedures both place the steps of depositing and patterning the resist before the metal deposition steps. This avoids the problems associated with undercut. In the second procedure, outlined in Figure 4.13, the resist is deposited and patterned, and the metal is then deposited via electroplating. ITO can be used as a plating base, but for an initial test 20 nm of Ti/Pt was used to seed the growth. Figure 4.14 shows a preliminary result of this technique, and also the primary disadvantage: poor quality of the metal layer. Although the gap size is again good, the high granularity of the metal would likely result in a poor quality antenna. The quality could be improved through the use of a different metal for plating, for nickel the layer quality was smoother and the plating process more simple. Even if this difficulty could be overcome, the uniform plating layer would still exist below the antenna. This

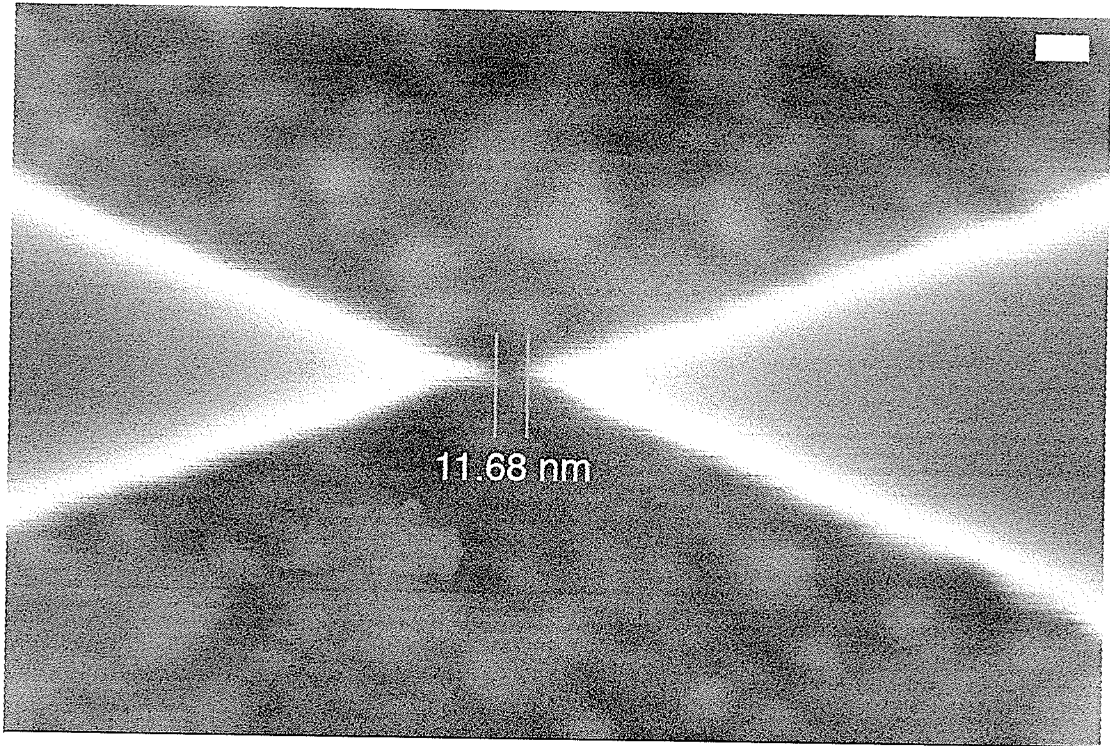


Figure 4.12: Exposure test of HSQ resist. Scale bar 20 nm.

would need to be etched, so this method is not likely to be feasible for our intentions.

The third procedure replaces the electroplating step with deposition and lift off. This is the method was the preferred EBL fabrication technique. The steps for creating the antennas are show in Figure 4.15. Beginning from a double PMMA resist roughly 90 nm thick that is spun on and exposed, roughly 5 nm of Cr was then sputtered for use as an adhesion layer followed by 30 nm Au also deposited via sputtering. Following liftoff 40 nm remained, but there was significant sidewall coverage. This adds a very thin edge around the antennas, the results of which are discussed in Section 4.15. The steps for creating the tapering channel and the window are shown in Figure 4.16. Two types of resist were used to make the lift-off capable of removing the large amount of metal for this stage. The negative resist was patterned with the electron beam, and the two metal layers were again sputtered and removed using liftoff. When the antennas were fabricated prior to the alignment features, there was less sidewall coverage than when the processes were reversed. Note that the bottom

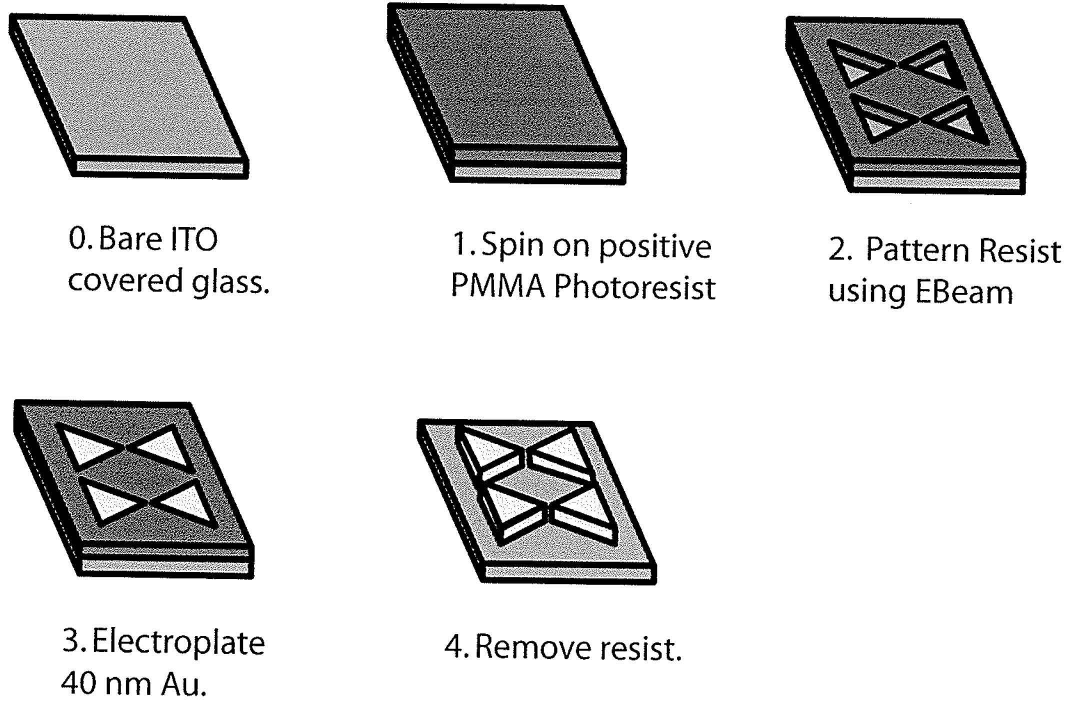


Figure 4.13: Second posible EBL fabrication technique.

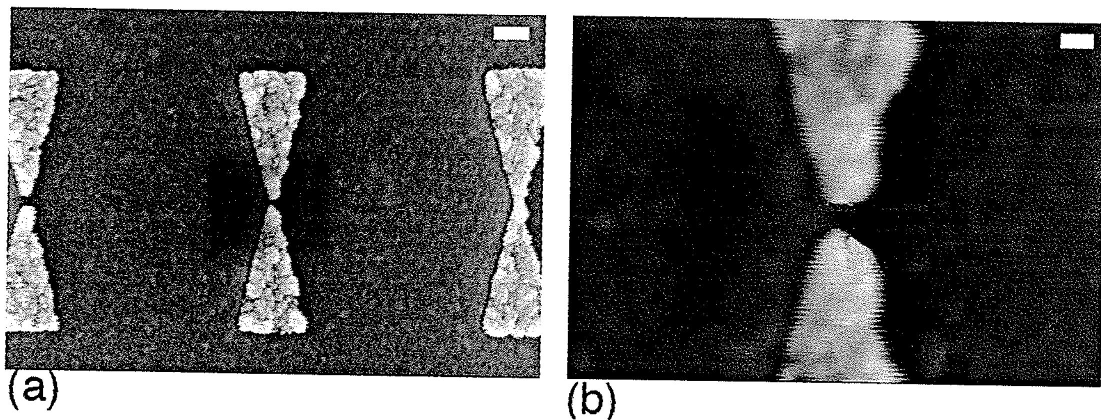
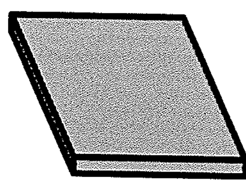
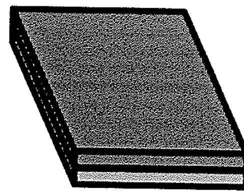
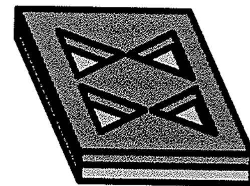
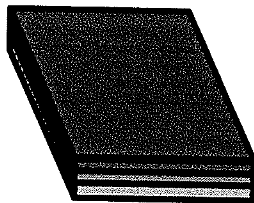


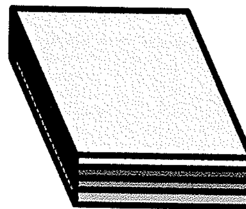
Figure 4.14: Nanoantennas fabricated using electroplating. Scale bars: (a) 100 nm (b) 20 nm











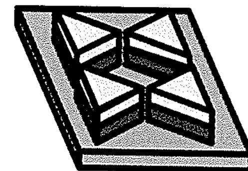
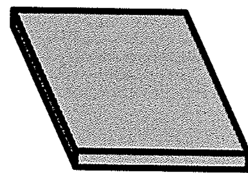
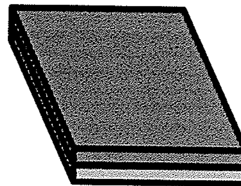
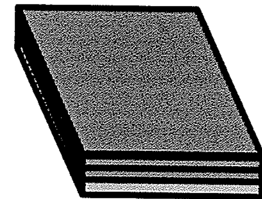
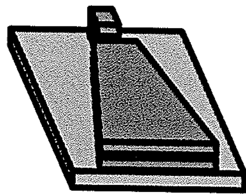


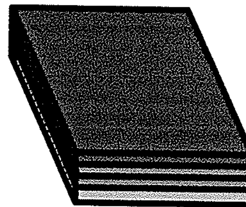
Figure 4.15: Final EBL fabrication technique for creating the antennas.

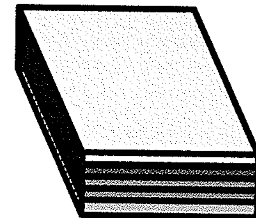












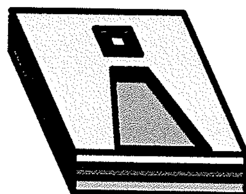


Figure 4.16: Final EBL fabrication technique for creating the tapering channels and windows.

ITO layer is not removed. To remove the ITO without damaging the antennas was deemed to difficult. Since ITO is a transparent conductor, its electrical properties at short wavelengths (i.e. visible and near-IR) should not significantly impact the behaviour of the antennas.

4.6.3 EBL Fabrication results

Figure 4.17 shows the results obtained via the lift-off technique. As can be seen, both the spacing uniformity and the gap size are better than the results obtained via focused ion beam fabrication. Whether or not this is due to the quality of the equipment, the accuracy of the calibration, or the limitations of techniques themselves is not known. Given the results obtained for this work however, EBL was the far superior technique.

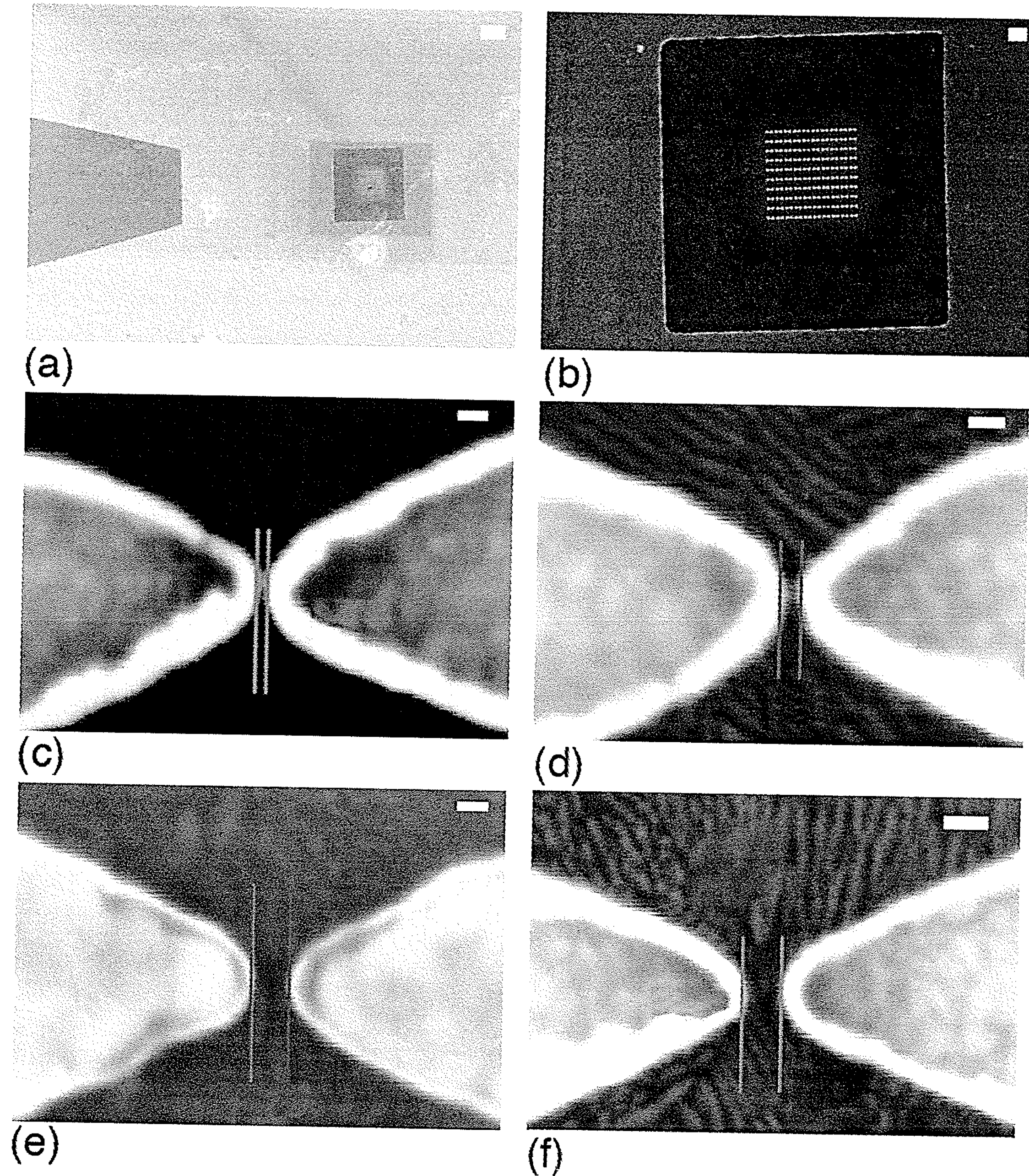


Figure 4.17: Results of electron beam fabrication. SEM images showing (a) Overview of window + tapering channel (b) close-up of window (c)-(f) close-up of individual antennas of varying gap sizes. Scale bar sizes: (a) 10 μm (b) 2 μm (c) 20 nm (d) 20 nm (e) 20 nm (f) 30 nm. Gap Sizes: (c) 7.03 nm (d) 11.68 nm (e) 23.60 nm (f) 26.99 nm

Chapter 5

Experiment

5.1 Introduction

To characterize the near-field around our nano-antennas is not an easy task. To attempt this we have selected two techniques that are likely to be capable of spatially imaging the optical near-field and be relatively wavelength independent, and definitely capable of detecting near-fields with λ=1550 nm. The techniques chosen attempt to detect optical near-fields though force measurement and through measurement of the scattered field. Both of these techniques utilize SPM systems.

5.2 Near-Field Detection via Force Measurement

5.2.1 Apparatus

Figure 5.1 shows the experimental setup for detection of the near-field of the antennas through force measurement. The antenna sits on top of the positioning apparatus described in Chapter 3. This is placed inside an Veeco[1] Enviroscope AFM which allows for the chamber pressure to be lowered using a vacuum pump. The Enviroscope, with the sample inside the chamber, is shown in Figure 5.2. The antenna is excited using an Agilent 81689A tunable laser. The laser is modulated at the resonant frequency of the cantilever in order to increase the sensitivity. This is done by selecting `Menu→Modulation`

[1]http://www.veeco.com/

source and choosing *External Analog*. The 5 V peak-to-peak signal is input to the front of the laser module. The cantilivered AFM probe scans over the surface four times per line (two times per mode): the first two scans are done in contact-mode and image the topography of the antennas, while the second two scans are done in lift-mode to detect the field components. A Stanford Research Systems SR830 DSP lock-in amplifier is used to extract the force component during the lift mode scan. The deflection signal can also be passed to a spectrum analyzer (not shown in the block diagram). The spectrum analyzer allows for a simple verification of whether the tip is sticking to the surface during lift mode. If there is a peak at the resonant frequency then the tip is not sticking to the surface.

The AFM is outfitted with a CSCS12 contact mode cantilever, purchased from Silicon-MDT, Ltd. (now part of μmasch[2]), coated with a W_2C layer to allow for electric field measurement. The CSCS12 chip contains 6 cantilevered probes, 3 on each side. The shorter side containing cantilevers A-C was used, since these are stiffer and thus less likely to stick to the surface during lift mode. The typical characteristics of these cantilevers as given in Table 5.1. Uncoated, the tip radius is below 10 nm, and the full tip cone angle is less than 20°.

Table 5.1: CSCS12 contact mode probe specifications.

Cantilever Type	Length (μm)	ω_r {range} (kHz)	ω_r {typical} (kHz)	Force constant {range} (N/m)	Force constant {typical} (N/m)
A	110	70-110	90	0.5-1.0	0.7
B	90	130-190	160	1.2-2.3	1.7
C	130	60-90	75	0.4-0.8	0.6
D	300	12-18	15	0.03-0.07	0.05
E	350	8-12	10	0.02-0.04	0.03
F	250	16-24	20	0.06-0.10	0.08

The chamber pressure is lowered to around 10^{-5} Torr to increase the sensitivity of the cantilever. The laser is modulated at the resonant frequency of the cantilever. This modulates the laser signal by 5% about the current power. We

[2]http://www.spmtips.com/

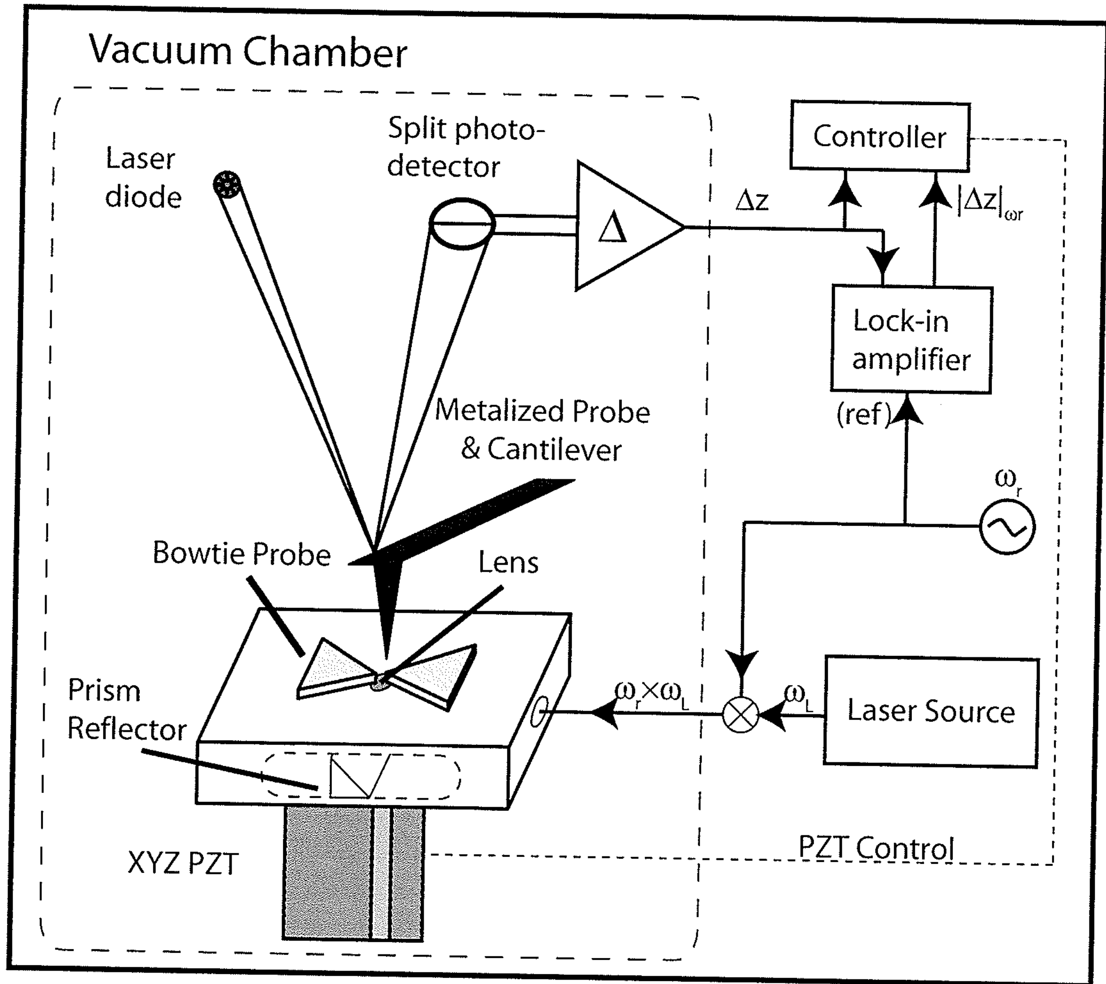


Figure 5.1: Block Diagram for detection though force measurement.

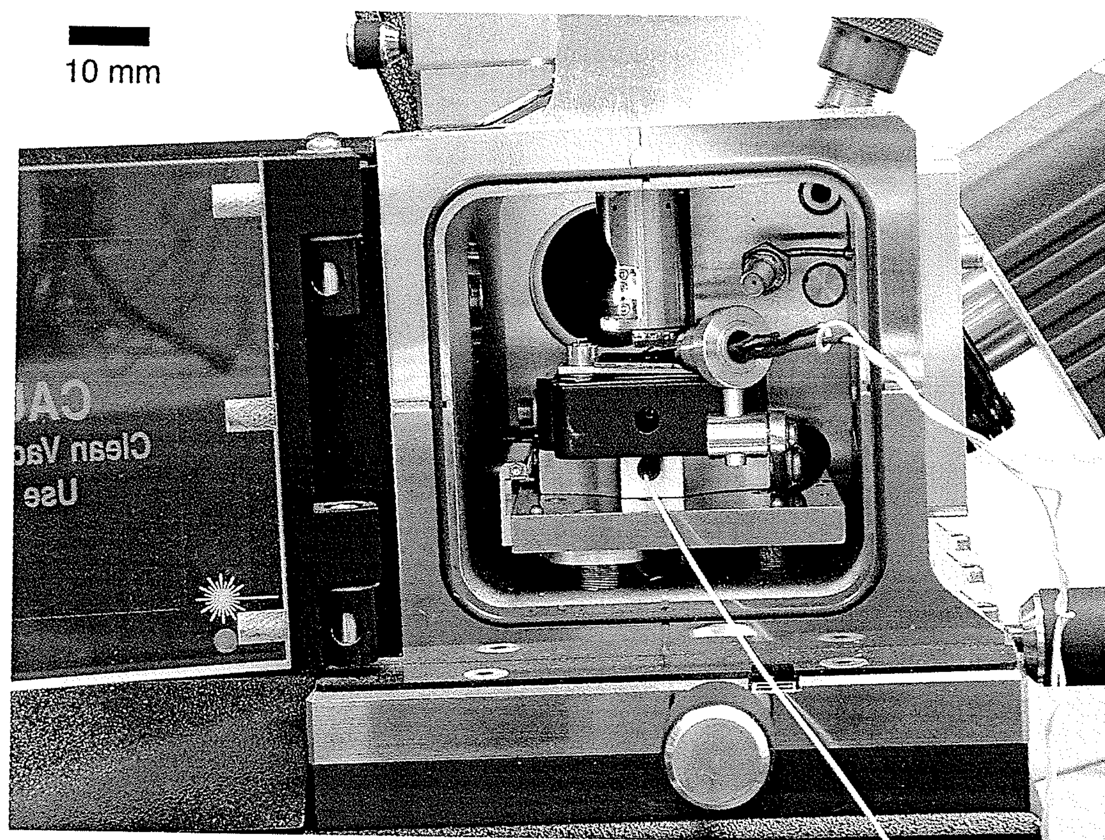


Figure 5.2: Digital Instruments Enviroscope.

can thus model the voltage on the tip as a

$$v_{tip} = v_{DC} + G_{tip} \times v_{laser} \left[1 + c_{modulation} \cos\left(\omega_r t\right)\right] \cos\left(\omega_L t\right) \tag{5.1}$$

and the voltage at the antenna as

$$v_{ant} = G_{ant} \times v_{laser} \left[1 + c_{modulation} \cos\left(\omega_r t\right)\right] \cos\left(\omega_L t\right), \tag{5.2}$$

where G_{ant} is near field gain of the antenna in the region where the tip is above, G_{tip} is the gain of the tip, and $c_{modulation} = 0.05^{1/2}$.

The force on the tip is proportional to

$$F \propto \left(v_{tip} - v_{antenna}\right)^2. \tag{5.3}$$

This results in the force on the cantilever being

$$F \propto F_{DC} + F_{\omega_r} + F_{2\omega_r} + F_{HF}, \tag{5.4}$$

where F_{HF} are all the force terms at frequencies above $2\omega_r$ (thus at of near multiples of ω_L). The forces terms at ω_r are:

$$F_{\omega_r} \propto c_{modulation} v_{laser}^2 \left[G_{ant}^2 + 2G_{ant}G_{tip} + G_{tip}^2\right]. \tag{5.5}$$

Assuming that $G_{ant} \gg G_{tip}$, at least in the gap, measuring the force at ω_r should give us a measurement of the near-field of the antenna. Note that the value of v_{DC} does not affect the signal at ω_r.

5.2.2 Results

Figures 5.3 show results of the force measurement on one of the probes fabricated with the FIB. In Figure 5.3 we see that there appears to be few effects from the laser turn on. Although the scales of the profiles have changed, the field distribution appears the same. This suggests that the topography of the antenna is the dominant contributor to the force measurement, and the near field effects are not significant 50 nm from the surface.

Figure 5.4 shows simulation results for the field along the x-axis 50 nm

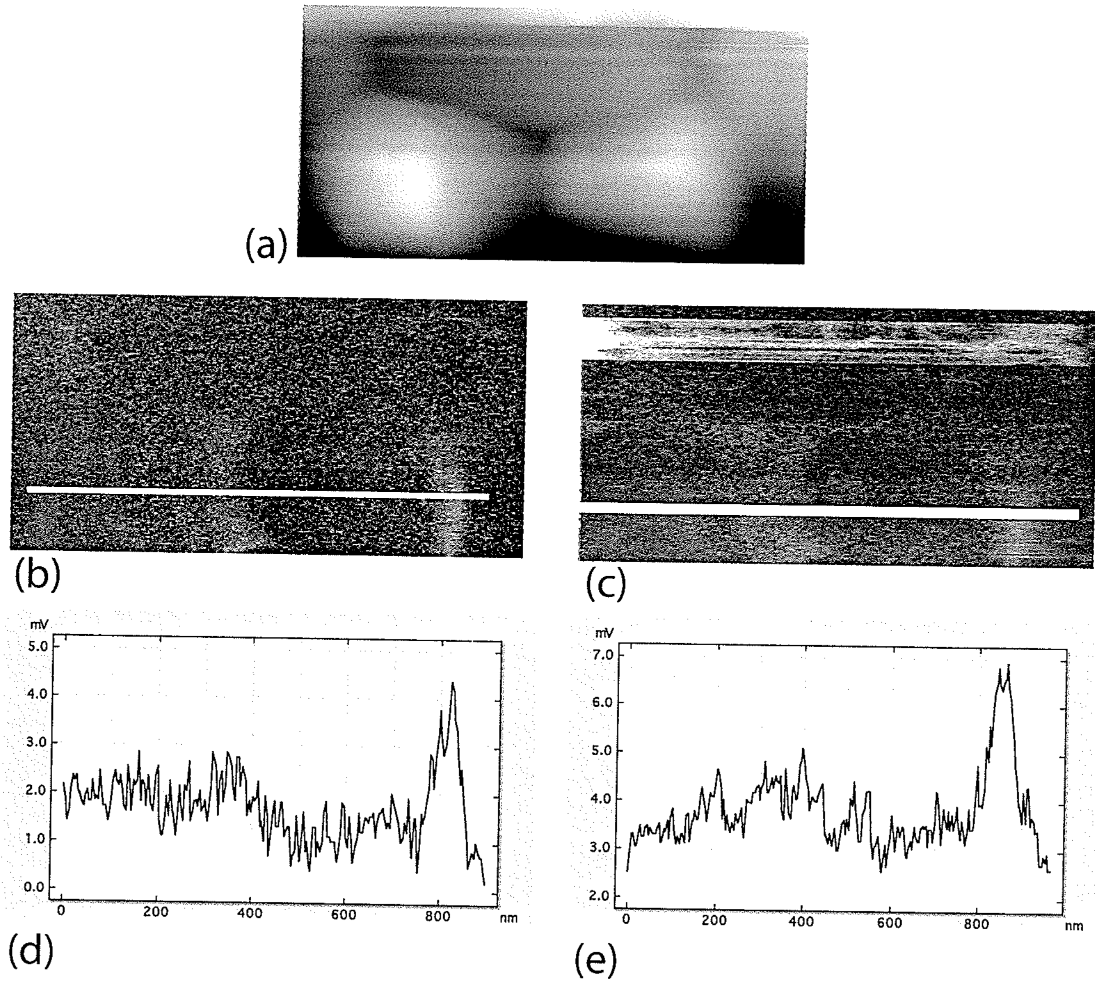


Figure 5.3: Nanoantenna force measurement results (all images 1 μm×500 nm). (a) topographic image, (b) force image at 50 nm lift with no laser power, (c) force image at 50 nm lift for P = 300 μW, (d) Average profile of outlined area in (b), Average profile of outlined area in (c).

above the surface for gap sizes of 10, 20 and 30 nm, while Figure 5.5 shows the simulation results along the y-axis. This shows that is it unlikely that we will be able to observe and evidence of amplification in the gap. Additionally, the effects of the different gap sizes, while very significant near the antenna, have little effect on the field 50 nm away.

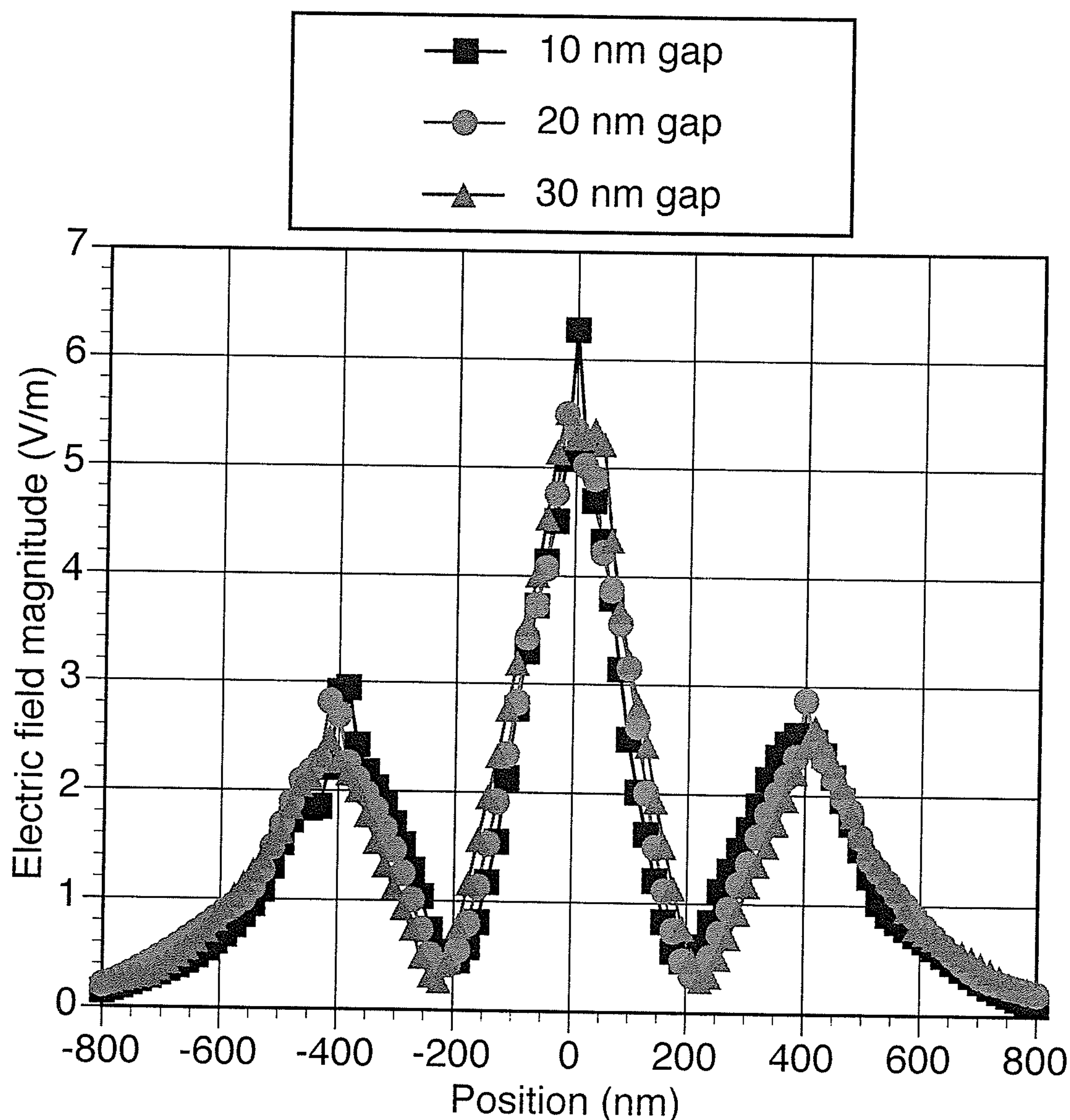


Figure 5.4: Simulation results of fields along the x-axis 50 nm above the antenna.

Small lift distance scans were attempted in order to try to better image the near-field, but whenever the lift-height was set below 50 nm the tip would

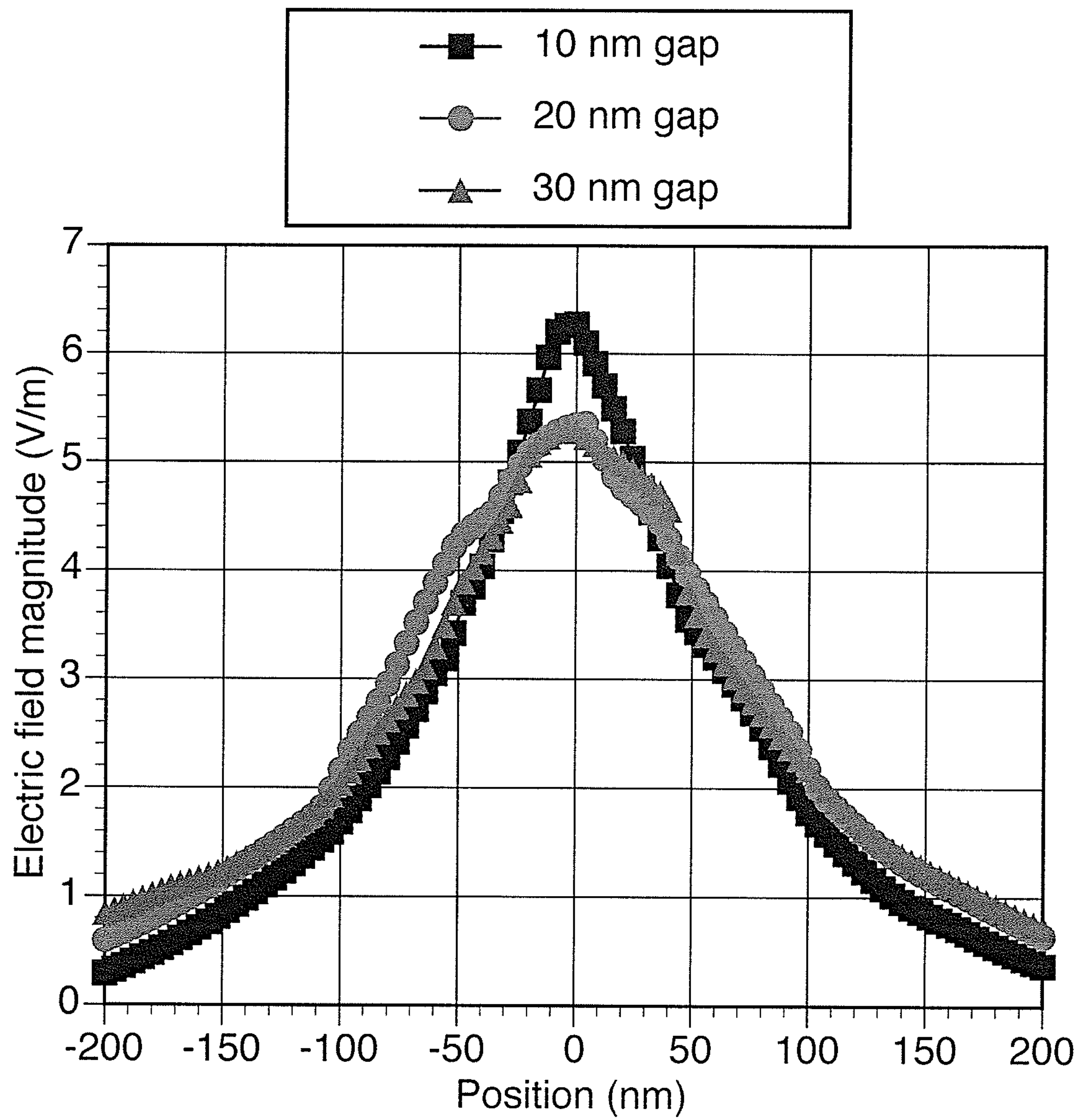


Figure 5.5: Simulation results of fields along the y-axis 50 nm above the antenna.

tend to stick to the surface. This was identified by the lack of a resonant peak during the lift stage of the scan. The use of stiffer cantilevers with a higher resonant frequency were tested, but did not give a noticeable improvement. Additionally, the determination of whether lift mode was working properly was significantly more complicated. Because of these difficulties, we did not attempt to image the EBL fabricated antennas.

5.3 Near-Field Detection via Scattered-Field Measurement

5.3.1 Apparatus

Taking in to account the difficulties with our force detection measurement, we devised a system to measure the near-field where lift mode was not employed. Figure 5.6 illustrates the apparatus used to image the near-field by measuring the scattering intensity off of a probe tip operating in tapping mode. The apparatus is very similar to the equipment used in apertureless NSOM, with the key difference being the lack of illumination source.

The antenna is illuminated in the the same fashion as the force measurement setup. This will again generate a strongly-localized near-field in the gap region of the antenna. A probe is driven at its resonant frequency and scanned over the surface. The detector described in section 3.5 is pointed towards the tip and collects the scattered light. The scattered field signal is passed to a Stanford Research Systems SR844 RF lock-in amplifier which extracts the signal component at the drive frequency of the cantilever. This signal is passed to the control to generate the near-field image. A similar apparatus was used to measure an optical near-field in [25].

This experiment used the point probe-plus non contact (ppp-nch) probe from nanoscience instruments[3]. The specifications for this probe is given in Table 5.2

[3] www.nanoscience.com

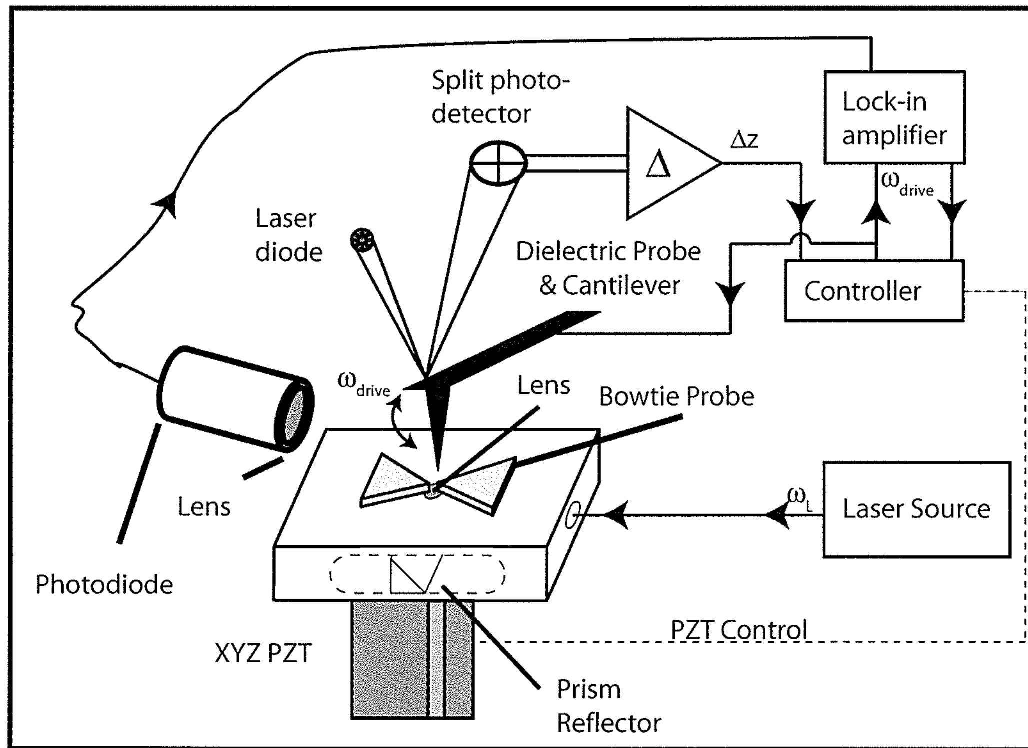


Figure 5.6: Block Diagram for detection though scattered field measurement.

Table 5.2: PPP-NCH tapping mode probe specifications.

Length (μm)	ω_r {range} (kHz)	ω_r {typical} (kHz)	Force constant {range} (N/m)	Force constant {typical} (N/m)	Tip radius {typical} (nm)
125	204-497	320	10-130	42	$\leq$7

5.3.2 Results

Figure 5.7 shows a typical scan of one of the EBL fabricated probes at different laser powers. Immediately we can see that scattered-field detection offers a much cleared image of the antennas, and the results are much more convincing than our force-detection measurements. When the laser was turned off, the scattered field signal was drastically reduced. There was still a small background signal however. Placing a filter in between the lens and the detector to remove all light with $\lambda \neq 1550$ nm would likely remove this signal.

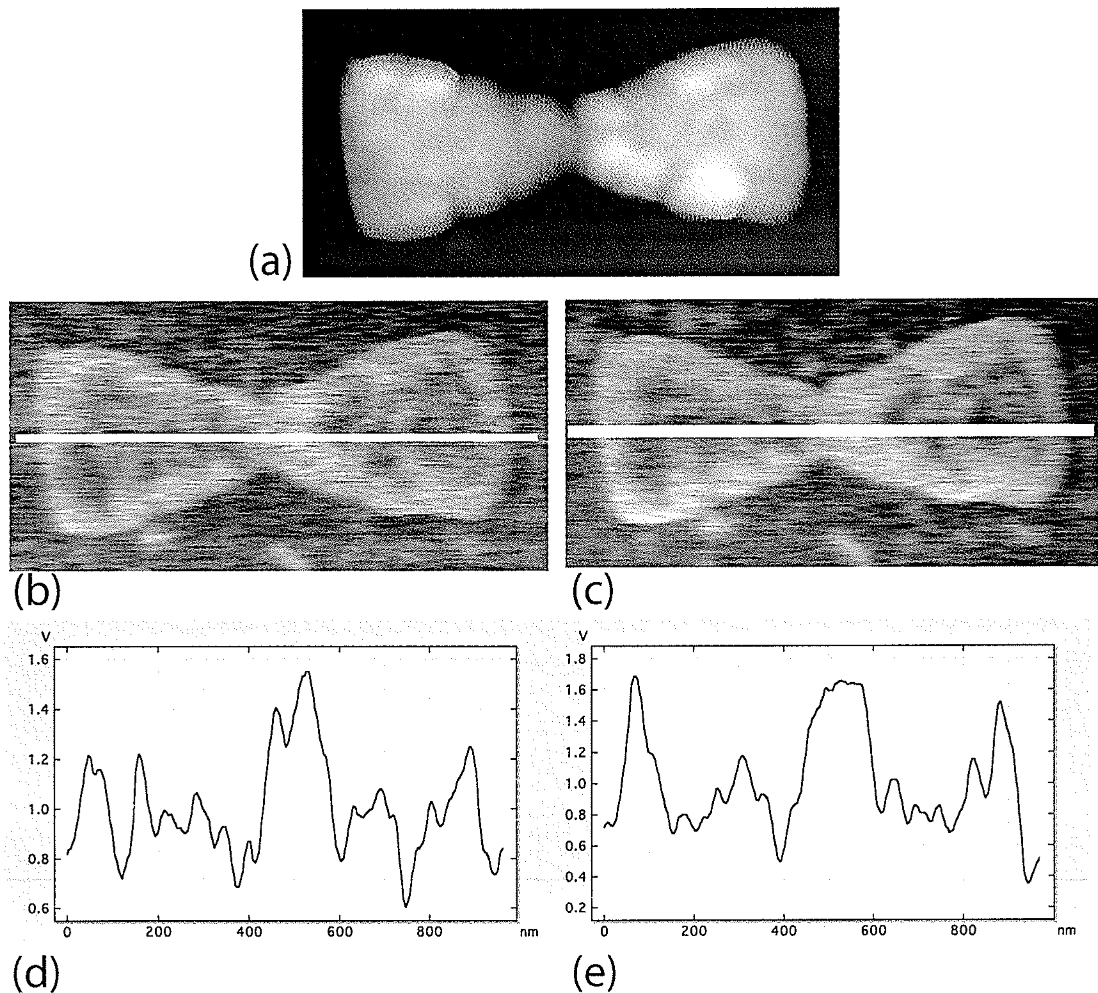


Figure 5.7: Nanoantenna scattered-field measurement results (all images 1 μm$\times$500 nm). (a) topographic image, (b) scattered-field image for P=500 μW, (c) scattered-field image for P = 750 μW, (d) Average profile of outlined area in (b), Average profile of outlined area in (c).

The first thing that is apparent about our measurements is that we are not observing the very high enhancement in the gap region that our simulation results predicted. We can still observe very definite enhancement occurring at the edges. This signal was present in both measurements from the FIB fabricated probes and the EBL fabricated probes, but was much stronger in the latter. This is thought to be due to field enhancement caused by the topographic defects on the antennas. This is likely higher in the EBL fabricated antennas due to the significant sidewall coverage. This unpredicted enhancement has not been documented in prior works on nanoantennas. All of the previous measurements on nanoantennas lacked the resolution to determine the source of the near-field enhancement, and we will investigate this further. The resolution of the AFM tip is not high enough to tell whether we have a well formed gap however, so a better comparison between the SEM images of the probes and the scattered field measurement is necessary before making any conclusions. There exists a possibility that the antennas imaged to date lack a well formed gap and are shorted at the bottom, which would preserve the enhancement at the edges but not have an enhancement factor in the gap.

Chapter 6

Conclusions

The limitations of conventional optics in imaging and materials processing have spurred investigation into near-field sources and techniques. Nanoscale optical antenna offer significant promise for use as strongly-localized high-powered sources, and have a wide range of potential applications such as imaging, localized spectroscopy, material processing, and magnetic storage[41]. Their design and preliminary characterization has already been explored at 458 nm and 830 nm. This work investigated the simulation, fabrication and characterization of nanoantennas operating at 1550 nm and gives the first image of the antenna near-fields.

The results of electromagnetic simulations showed the importance of the total antenna length and the gap size between the two halves of the probe. When the antenna length is varied between 750 nm and 6 μm, the field enhancement varies between 140× to as low as 5×. For gap sizes above 20 nm, the enhancement factor drops to below 50×, and continues to decrease for large gaps. The effects of the flare angle, the antenna thickness, and the metal used to fabricate the probe were also calculated and found to be less significant, though still important. Varying the metal thickness between 20 nm and 100 nm varies the enhancement factor by 50%, and varying the flare angle between 20° and 90° varies it by 20%.

An apparatus was built to focus and position a collimated optical beam on to the antenna inside a SPM unit. When combined with specially designed alignment features on the sample, this proves to be a highly compact and relatively simple system that gives the precision required to focus a 30 μm spot

over a window containing a nanoantenna array.

Both focused ion beam milling and electron beam lithography were explored as fabrication techniques. Of these methods, electron beam lithography offered better results due to its low drift and smaller gap sizes. Additionally, all steps of the process can be completed using one piece of equipment, whereas ion milling must be combined with optical lithography in order to process the necessary alignment features.

Two analysis techniques were explored to characterize these antennas. Near-field imaging through force detection was not successful in characterizing the antenna. Although a signal from the excitation source could be measured, the dominant signal was dependent on the topographic data. Additionally, the lift-mode scan would not function properly any closer than 50 nm, which is outside the region of high field enhancement.

Near-field imaging through scattered-field measurements is a much more promising technique. Preliminary measurements give a definite subwavelength resolution, and the signal is very independent of topographic data. The high enhancement seen at the edges is most likely due to the sidewalls that remained following liftoff. The field in the gap region is higher, but is not at the levels predicted by the simulations. We have observed a previously undocumented but significant enhancement at the edges of the antennas that is likely due to topographic defects. Further testing is necessary to fully characterize this effect and the enhancement factor in the gap.

Appendix A

Fourier optics and optical limits

While the diffraction limit arises from the far-field assumption and is typically the starting point for any analysis of sub-wavelength imaging, the derivation of the limit is rarely stated. It is instructive to understand the derivation, and this as well as the basic analysis techniques are shown in this chapter for the convenience of the interested reader. Most of this information can be found in any standard optics textbook. The general progression of the derivation has been taken from [38].

A.1 Fourier Optics

To determine the resolution limit in the far-field, it is necessary to employ the principles of Fourier optics. We begin with the concept that an arbitrary function of two variables, $f(x, y)$, can be expressed as a superposition of harmonic functions of the form $F(v_x, v_y) \exp\left[-j2\pi(v_x + v_y)\right]$, where $F(v_x, v_y)$ is the complex amplitude and v_x and v_y are the spatial frequencies in the x & y directions. Additionally, we note that in an arbitrary plane, $U(x, y, z)$ is a spatial harmonic function. Specifically, in the $z = 0$ plane, if the spatial harmonic frequency is less than the inverse wavelength $1/\lambda$ then there exists a one-to-one correspondence between $U(x, y, z)$ and the function $f(x, y) = U(x, y, 0)$[38]. Combining these two ideas gives us the basic principle of Fourier optics. We can analyze an arbitrary traveling wave as a superposition of plane waves, thus by knowing how a system modifies incoming plane waves, we can determine the output for any incident signal.

A.2 Light in Free space

The analysis of the propagation of a monochromatic plane wave in free space is the first step to determining the resolution limit of far-field optics. We consider the propagation of a plane wave in free space with wavelength λ between xy planes located at $z = 0$ and $z = d$. We thus want to calculate complex amplitude at the output $g(x, y) = U(x, y, d)$ given the input amplitude $f(x, y) = U(x, y, 0)$.

The wave must satisfy the Helmholtz equation. Since this equation is linear, the system that transforms $f(x, y)$ in to $g(x, y)$ (See Figure A.1) must be linear as well. Additionally, free space is invariant to shifts in the coordinate system, thus the system must possess this property (time invariance) as well. A linear, time-invariant (LTI) system can be characterized by either its impulse response $h(x, y)$, or by its transfer function $H(v_x, v_y)$. We will derive the latter. We begin with the plane wave $U(x, y, z) = A \exp[-j(k_x x + k_y y + k_z z)]$, where

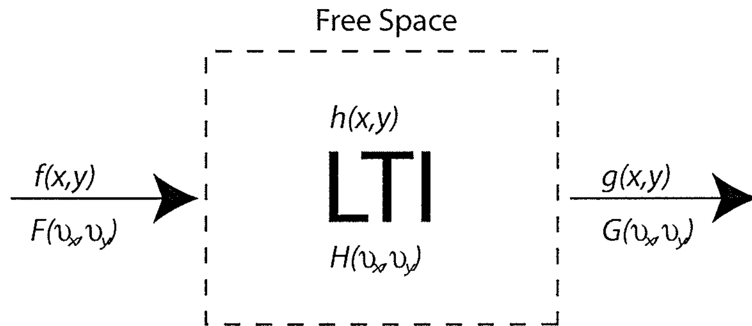


Figure A.1: Free space transfer function.

$k_x = 2\pi v_x$, $k_y = 2\pi v_y$ and $k^2 = k_x^2 + k_y^2 + k_z^2 = 1/\lambda^2$. As shown earlier, we have a correspondence in the $z = 0$ plane with $f(x, y) = A \exp[-j(k_x x + k_y y)]$. In the $z = d$ plane the output wave, we have $g(x, y) = A \exp[-j(k_x x + k_y y + k_z d)]$. Writing the z-component of the wave-vector as

$$k_z = \left(k^2 - k_x^2 - k_y^2\right)^{1/2} = 2\pi \left(\frac{1}{\lambda^2} - v_x^2 - v_y^2\right)^{1/2}, \tag{A.1}$$

we can then determine the transfer function of free space as:

$$H(\upsilon_x, \upsilon_y) = \frac{g(x,y)}{f(x,y)} = \exp\left(-jk_z d\right) = \exp\left[-j2\pi\left(\frac{1}{\lambda^2} - \upsilon_x^2 - \upsilon_y^2\right)^{1/2} d\right]. \quad \text{(A.2)}$$

From this result we can observe an important aspect of the propagation of optical waves. If the sum of the spatial frequencies, $\upsilon_x^2 + \upsilon_y^2$ is less than or equal the inverse of the wavelength squared, $1/\lambda^2$, the magnitude of H, $|H(\upsilon_x, \upsilon_y)|$ will be unitary, the phase will be a function of the spatial frequencies, and the plane wave will propagate as usual. If $\upsilon_x^2 + \upsilon_y^2 > 1/\lambda^2$ however, the square root term becomes imaginary, which causes the entire exponential function to be real-valued. This results in an attenuation of the plane wave with distance. From this result we can clearly see that *any features with spatial frequencies greater than $1/\lambda$ will not propagate to the far field.*

A.3 Approximation methods

A.3.1 The Fresnel Approximation

There are two different approximations that can be made for the propagation of light in free-space, each of which leads to a different analysis of the maximum theoretical resolution. The simplest assumption is that spatial frequencies of $f(x,y)$, υ_x and υ_y are such that $\upsilon_x^2 + \upsilon_y^2 \ll 1/\lambda^2$. Since this leads to the result that the angle of the wavevector, $\theta_x = \sin^{-1}(\lambda\upsilon_x)$ and $\theta_y = \sin^{-1}(\lambda\upsilon_y)$ are small, these can be approximated as $\theta_x \approx (\lambda\upsilon_x)$ and $\theta_y \approx \lambda\upsilon_y$. Writing $\theta^2 = \theta_x^2 + \theta_y^2 \approx \lambda^2(\upsilon_x^2 + \upsilon_y^2)$ allows us to approximate the square root term in (A.2) as

$$\left(\frac{1}{\lambda^2} - \upsilon_x^2 - \upsilon_y^2\right)^{1/2} \approx \frac{1}{\lambda}(1-\theta^2)^{1/2} = \frac{1}{\lambda}\left(1 - \frac{\theta^2}{2} + \frac{\theta^4}{8} - \ldots\right) \quad \text{(A.3)}$$

By neglecting term of order higher than 2 , we obtain what is know as the *Fresnel approximation* for the free space transfer function

$$H(\upsilon_x, \upsilon_y) \approx \exp\left(-jkd\right)\exp\left[j\pi\lambda d(\upsilon_x^2 + \upsilon_y^2)\right] \quad \text{(A.4)}$$

For this approximation to be valid, we require the third term on the right side of (A.3) to be much less than π (the maximum possible value for a forward propagating wave) for all values of θ. This means

$$\frac{2\pi d}{\lambda}\left(\frac{\theta^4}{8}\right) \ll \pi \Rightarrow \frac{\theta^4 d}{4\lambda} \ll 1, \tag{A.5}$$

Now if the largest angle is θ_m, and the largest radial distance in the plane $z = d$ is a, then $\theta_m \approx a/d$ and the condition for the validity of the Fresnel approximation becomes

$$\frac{\theta^4 d}{4\lambda} \approx \frac{d}{4\lambda}\left(\frac{a^4}{d^4}\right) = \frac{1}{4}\left(\frac{a^2}{\lambda d}\right)\left(\frac{a^2}{d^2}\right) \approx \frac{N_F \theta_m^2}{4} \ll 1, \tag{A.6}$$

where $N_F = a^2/\lambda d$ is known as the Fresnel number.

To apply the Fresnel approximation to calculating $g(x, y)$, we start with the Fourier transform of $f(x, y)$

$$F(v_x, v_y) = \iint f(x, y) \exp\left[j2\pi(v_x x + v_y y)\right] dx\, dy. \tag{A.7}$$

The product $F(v_x, v_y)H(v_x, v_y)$ gives the frequency domain representation of the output plane wave. We then perform an inverse Fourier transform to obtain the representation in the output plane

$$g(x, y) = \iint F(v_x, v_y)H(v_x, v_y) \exp\left[-j2\pi(v_x x + v_y y)\right] dv_x\, dv_y. \tag{A.8}$$

By applying the Fresnel approximation for $H(v_x, v_y)$, this reduces to

$$\boxed{g(x, y) = \exp(-jkd) \iint F(v_x, v_y) \exp\left[-2\pi\lambda d(v_x^2 + v_y^2)\right] \exp\left[-j2\pi(v_x x + v_y y)\right] dv_x\, dv_y}. \tag{A.9}$$

The output can be determined by performing our calculations in the spatial domain as well. We begin by taking the inverse Fourier transform of (A.4), which is

$$h(x, y) \approx h_0 \exp\left(-jk\frac{x^2 + y^2}{2d}\right), \tag{A.10}$$

where $h_0 = (j/\lambda d)\exp(-jkd)$. Since multiplication in the frequency domain

corresponds to convolution in the spatial domain the output is

$$g(x,y) = \iint f(x',y')h(x-x',y-y')\,dx'\,dy' \tag{A.11}$$

By applying the Fresnel approximation (A.10) this becomes

$$\boxed{g(x,y) = h_0 \iint f(x',y') \exp\left[-j\pi \frac{(x-x')^2+(y-y')^2}{\lambda d}\right] dx'dy'}, \tag{A.12}$$

A.3.2 The Fraunhofer approximation

Our derivation of the Fraunhofer approximation begins with the Fresnel approximation of the convolution relation between $f(x,y)$ and $g(x,y)$, given in (A.12). If $f(x,y)$ exists only inside a small area of radius b, and d is large enough that

$$N_F' = \frac{b^2}{\lambda d} \ll 1 \tag{A.13}$$

then we can expand the argument inside the square root in (A.12) as follows:

$$\begin{aligned} \frac{-j\pi}{\lambda d}[(x-x')^2+(y-y')^2] &= \frac{-j\pi}{\lambda d}[(x^2+y^2)+(x'^2+y'^2)-2(xx'+yy')] \\ &\approx \frac{-j\pi}{\lambda d}[(x^2+y^2)-2(xx'+yy')] \end{aligned} \tag{A.14}$$

since $f(x,y) \neq 0$ only for $x'^2+y'^2 \leq b^2$ and $b^2\lambda d$ is very small. Substituting into (A.12) and bringing the non-primed terms outside the integral gives

$$g(x,y) = h_0 \exp\left(-j\pi\frac{x^2+y^2}{\lambda d}\right) \iint f(x',y') \exp\left(j2\pi\frac{xx'+yy'}{\lambda d}\right) dx'\,dy'. \tag{A.15}$$

By regarding $x/\lambda d$ and $y/\lambda d$ as the spatial frequencies v_x and v_y, respectively, we can regard the integral term in (A.15) as the Fourier transform of f(x,y) at these spatial frequencies. If the points in the $z = d$ plane are also confined to a small circle of radius a such that $N_F \ll 1$ then we can also disregard the first exponential and write

$$\boxed{g(x,y) \approx h_0 F\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right)}. \tag{A.16}$$

This is known as the *Fraunhofer approximation*, and requires both Fresnel

numbers N_F and N'_F to be small.

A.4 Diffraction

We will now examine the diffraction of a monochromatic beam of light though an aperture. For the ray approximation of light, the diffraction pattern would simply be a shadow of the aperture. The wave nature of light means that the diffraction pattern will contain features from the interaction of the waves with each other, and this changes based on the aperture shape, the incident wavelength, and the distance between the aperture and observation planes. While the exact nature interaction between the aperture and the incident wave is a difficult problem, it can be simplified through the use of the approximations developed in the previous section.

To construct a theory of diffraction based on these approximations, we make the assumption that the transmitted wave immediately following the aperture is nonzero only in the non-opaque regions. This can be expressed mathematically as

$$f(x, y) = U(x, y)p(x, y). \tag{A.17}$$

where

$$p(x, y) = \begin{cases} 1 & \text{inside the aperture} \\ 0 & \text{outside the aperture} \end{cases} \tag{A.18}$$

$p(x, y)$ is known as the *aperture function*. From $f(x, y)$ we can then determine the complex amplitude that would be projected on to an observation plane a distance d away. The measured diffraction pattern $I(x, y) = g(x, y)$ is referred to as the *Fraunhofer diffraction* or *Fresnel diffraction*, depending on the approximation made for the free space propagation following the aperture. The former is what is assumed when the far-field assumption is discussed, and is what is responsible for the Abbé diffraction limit.

It should be noted that (A.18) is not exact. Beyond the aperture the wave propagates in all directions and will reach points on the back of the screen. To obtain an exact answer to the diffraction problem is a difficult task that requires an exact solution to the Helmholtz equation with the boundary conditions given by the aperture.

A.4.1 Fraunhofer Diffraction

In Fraunhofer diffraction we make the assumption that the incident wave is multiplied by the aperture function and that the propagation beyond the aperture is modeled by the Fraunhofer approximation. As we have seen, for this approximation to be valid it is required that the distance d that separates the aperture and the projection is large enough that $N_F' = b^2/\lambda d \ll 1$, where b is the largest radial distance of the aperture.

If the incident wave is a plane wave with uniform intensity I_i propagating in the z direction, then $U(x,y) = I_i^{1/2}$ and $f(x,y) = I_i^{1/2} p(x,y)$. The Fraunhofer approximation thus gives

$$g(x,y) \approx I_i^{1/2} h_0 P\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right) \tag{A.19}$$

The rightmost term, $P(v_x, v_y)$ is the Fourier transform of $p(x,y)$ at spatial frequencies $v_x = x/\lambda d$, $v_y = y/\lambda d$ given in (A.7):

$$P(v_x, v_y) = \iint p(x,y) \exp\left[j2\pi(v_x x + v_y y)\right] dx dy. \tag{A.20}$$

Using (A.19), the Fraunhofer diffraction is thus

$$\boxed{I(x,y) = |g(x,y)|^2 = \frac{I_i}{(\lambda d)^2} \left| P\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right)\right|^2} \tag{A.21}$$

We now calculate the Fraunhofer diffraction from a circular aperture, as shown in Figure A.2. We begin by calculating the Fourier transform of the circular aperture. Starting from (A.20), we convert to cylindrical coordinates, using $x = \rho \cos\phi$, $y = \rho \sin\phi$, $x' = q\cos\Phi$ and $y' = q \sin\Phi$ (with $v_x = x'/\lambda d$, $v_y = y'/\lambda d$):

$$\begin{aligned} P(v_x, v_y) &= \int_0^{D/2} \int_0^{2\pi} \exp\left[j\left(\frac{2\pi\rho q}{\lambda d}\right)(\cos\phi\cos\Phi + \sin\phi\sin\Phi)\right] \rho \, d\rho \, d\phi \\ &= \int_0^{D/2} \int_0^{2\pi} \exp\left[j\left(\frac{2\pi\rho q}{\lambda d}\right)\cos(\phi - \Phi)\right] q \, dq \, d\phi. \end{aligned} \tag{A.22}$$

From symmetry we know that the result will not depend on Φ, so we set $\Phi = 0$.

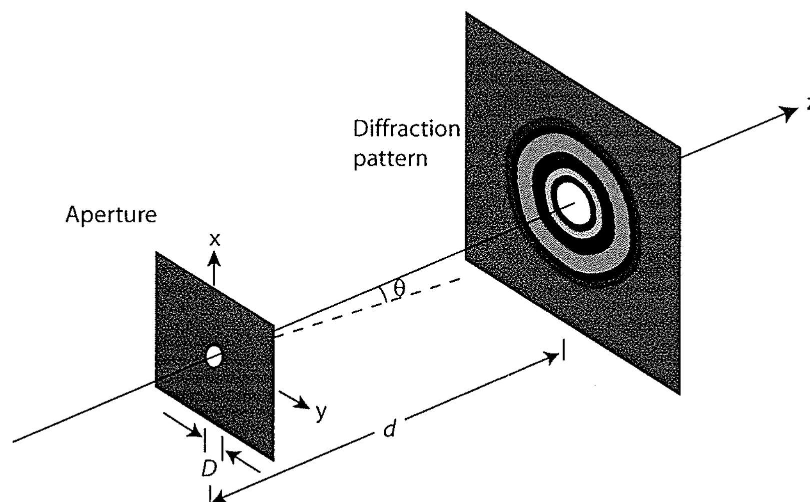


Figure A.2: Fraunhofer diffraction from a circular aperture.

We now use the function

$$J_0(u) = \frac{1}{2\pi} \int_0^{2\pi} e^{ju\cos v}\, dv, \tag{A.23}$$

known as a *Bessel function of the first kind*. By plugging this in to (A.22), the expression transforms to

$$P(v_x, v_y) = 2\pi \int_0^{D/2} J_0\left(\frac{2\pi\rho q}{\lambda d}\right) q\, dq. \tag{A.24}$$

We now make use of the recurrence relationship for Bessel functions

$$\int_0^u u' J_0(u')\, du' = uJ_1(u). \tag{A.25}$$

We let $u' = 2\pi\rho q/\lambda d$, so

$$\frac{du'}{dq} = \frac{2\pi\rho}{\lambda d} \Rightarrow qdq = \frac{\lambda dq}{2\pi\rho} \tag{A.26}$$

This changes the upper bound on the integral to $u = \frac{\pi D\rho}{\lambda d}$ and the expression

becomes

$$
\begin{aligned}
2\pi \int_0^{\frac{\pi D\rho}{\lambda d}} \frac{\lambda dq}{2\pi\rho} J_0(u')\, du' &= 2\pi \int_0^{\frac{\pi D\rho}{\lambda d}} \frac{\lambda^2 d^2}{4\pi^2\rho^2} u' J_0(u')\, du' \\
&= \frac{\pi D}{4u^2} 2 \int_0^u u' J_0(u')\, du'.
\end{aligned} \tag{A.27}
$$

Now, using the recurrence relation this becomes

$$
P(v_x, v_y) = \frac{\pi D^2}{4} \frac{2J_1(u)}{u} \tag{A.28}
$$

Expanding u and substituting everything in to A.21 gives

$$
\boxed{I(x,y) = I_0 \left[\frac{2J_1(\pi D\rho/\lambda d)}{\pi D\rho/\lambda d} \right]^2,} \tag{A.29}
$$

where

$$
\rho = (x^2 + y^2)^{1/2} \tag{A.30}
$$

and

$$
I_0 = \left(\frac{\pi D^2}{4\lambda d} \right)^2. \tag{A.31}
$$

A.4.2 The Fraunhofer diffraction limit

Equation (A.29) enables us to determine the resolution limit of a system where Fraunhofer diffraction is assumed. We assume now that our observation plane is now at $z = 0$, so we are viewing a sample in plane $z = d$ through a lens of diameter D. While we derived the diffraction limit under the assumption that the medium between the two planes was a vacuum, we can apply (A.29) to an arbitrary medium by replacing λ with λ/n, where n is the new index of refraction. We now introduce the concept of the *numerical aperture* of a lens. Numerical aperture characterizes the range of angles over which a lens can emit or absorb light. It is defined as

$$
NA = n \sin\theta, \tag{A.32}
$$

where n is the index of refraction and θ is the half-angle of the maximum cone of light that can enter or exit the lens. In the Fraunhofer limit, we thus have $\sin\theta = (D/2)/d$. We can thus rewrite the term inside the Bessel function in (A.29) (in an arbitrary medium) as

$$\frac{\pi D\rho}{(\lambda/n)d} = \frac{2\pi n\sin\theta\rho}{\lambda} = \frac{2\pi NA\rho}{\lambda} \tag{A.33}$$

The first zero of $J_1(x)$ occurs at $x = 3.832$, so we find the radius of the zero of the Airy disk:

$$\frac{2\pi NA\rho}{\lambda} = 3.832 \Rightarrow \boxed{\rho = \frac{0.61\lambda}{NA}} \tag{A.34}$$

By applying the *Rayleigh criterion*, which states that two objects are distinguishable if the second object is located further than the first zero of the Airy disk of the first, we see that (A.34) gives us the diffraction limit in the far-field assumption.

Bibliography

[1] M. Abbé, Archiv. Mikroscop. Anat. Entwicklungsmech. **9** (1873).

[2] E. X. Jin and X. F. Xu, APPLIED PHYSICS LETTERS **88**, 153110 (2006).

[3] P. J. Schuck, D. P. Fromm, A. Sundaramurthy, G. S. Kino, and W. E. Moerner, PHYSICAL REVIEW LETTERS **94**, 017402 (2005).

[4] P. Muhlschlegel, H. J. Eisler, O. J. F. Martin, B. Hecht, and D. W. Pohl, SCIENCE **308**, 1607 (2005).

[5] B. Hecht *et al.*, JOURNAL OF CHEMICAL PHYSICS **112**, 7761 (2000).

[6] E. Synge, Philos. Mag. **6** (1928).

[7] D. W. Pohl, W. Denk, and M. Lanz, APPLIED PHYSICS LETTERS **44**, 651 (1984).

[8] U. Durig, D. W. Pohl, and F. Rohner, JOURNAL OF APPLIED PHYSICS **59**, 3318 (1986).

[9] A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, ULTRAMICROSCOPY **13**, 227 (1984).

[10] A. Harootunian, E. Betzig, A. Muray, A. Lewis, and M. Isaacson, JOURNAL OF THE OPTICAL SOCIETY OF AMERICA A-OPTICS IMAGE SCIENCE AND VISION **1**, 1293 (1984).

[11] E. Betzig, A. Harootunian, A. Lewis, and M. Isaacson, BIOPHYSICAL JOURNAL **47**, A407 (1985).

[12] E. Betzig, A. Lewis, A. Harootunian, M. Isaacson, and E. Kratschmer, BIOPHYSICAL JOURNAL **49**, 269 (1986).

[13] A. Lazarev, N. Fang, Q. Luo, and X. Zhang, Review Of Scientific Instruments **74**, 3679 (2003).

[14] D. Molenda, G. C. des Francs, U. C. Fischer, N. Rau, and A. Naber, OPTICS EXPRESS **13**, 10688 (2005).

[15] K. Lieberman, S. Harush, A. Lewis, and R. Kopelman, SCIENCE **247**, 59 (1990).

[16] N. Hosaka and T. Saiki, JOURNAL OF MICROSCOPY-OXFORD **202**, 362 (2001).

[17] W. X. Sun and Z. X. Shen, JOURNAL OF RAMAN SPECTROSCOPY **34**, 668 (2003).

[18] F. Zenhausern, M. O'Boyle, and H. K. Wickranmasinnghe, Applied Physics Letters **65**, 1623 (1994).

[19] M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge University Press, 1999).

[20] M. Inoue and K. Ohtaka, JOURNAL OF THE PHYSICAL SOCIETY OF JAPAN **52**, 3853 (1983).

[21] P. Barber and S. Hill, *Light Scattering by Particles: Computational Methods* (Wold Scientific, 1990).

[22] T. Kalkbrenner, M. Ramstein, J. Mlynek, and V. Sandoghdar, JOURNAL OF MICROSCOPY-OXFORD **202**, 72 (2001).

[23] T. J. Silva, S. Schultz, and D. Weller, APPLIED PHYSICS LETTERS **65**, 658 (1994).

[24] J. T. Krug, E. J. Sanchez, and X. S. Xie, JOURNAL OF CHEMICAL PHYSICS **116**, 10895 (2002).

[25] R. Quidant *et al.*, OPTICS EXPRESS **12**, 282 (2004).

[26] V. Lulevich and W. A. Ducker, APPLIED PHYSICS LETTERS **87**, 214107 (2005).

[27] F. Formanek, Y. De Wilde, and L. Aigouy, ULTRAMICROSCOPY **103**, 133 (2005).

[28] D. Courjon, K. Sarayeddine, and M. Spajer, OPTICS COMMUNICATIONS **71**, 23 (1989).

[29] R. C. Reddick, R. J. Warmack, and T. L. Ferrel, PHYSICAL REVIEW B **39**, 767 (1989).

[30] S. Takahashi, T. Fujimoto, K. Kato, and I. Kojima, NANOTECHNOLOGY **8**, A54 (1997).

[31] J. R. Sambles, G. W. Bradbery, and F. Z. Yang, CONTEMPORARY PHYSICS **32**, 173 (1991).

[32] R. D. Grober, R. J. Schoelkopf, and D. E. Prober, APPLIED PHYSICS LETTERS **70**, 1354 (1997).

[33] B. Rosner, IEEE transactions on antennas and propagation **50**, 670 (2002), 0018-926X.

[34] D. Palik, editor, *The Handbook of Optical Constants of Solids* (Academic, 1985).

[35] D. P. Fromm, A. Sundaramurthy, P. J. Schuck, G. Kino, and W. E. Moerner, Nano Letters **4**, 957 (2004).

[36] Y. D. Qin, D. L. Wang, S. F. Wang, and Q. H. Gong, CHINESE PHYSICS LETTERS **18**, 390 (2001).

[37] P. A. Champert *et al.*, OPTICS EXPRESS **12**, 4366 (2004).

[38] B. Saleh and M. C. Teich, *Fundamentals of Photonics* (Wiley-Interscience, 1991).

[39] S. Reyntjens and R. Puers, JOURNAL OF MICROMECHANICS AND MICROENGINEERING **11**, 287 (2001).

[40] M. Gad-el Hak, *The MEMS Handbook* (CRC Press, 2002).

[41] K. Sendur and W. Challener, Journal Of Microscopy-Oxford **210**, 279 (2003), Part 3.