Integral Equations Applied to Wave Propagation in Two-Dimensions: Modeling the Tip of a Near-Field Scanning Optical Microscope

September 13, 2000

C. M. Kelso[†],

P. David Flammer[†],

J. A. DeSanto[*],

R.T. Collins[†].

[†] Department of Physics,

[*] Department of Mathematical and Computer Sciences;

Colorado School of Mines, Golden CO 80401

Abstract

This paper presents a Green's function/Green's Theorem integral equation approach to numerically modeling two-dimensional, s polarized, wave propagation problems effectively for a variety of geometries. The model accurately calculates both near-fields and far-fields due to the minimal assumptions made on the behavior of the scattered radiation. The method was applied to modeling light emission from a Near-field Scanning Optical Microscope (NSOM) fiber tip. Several convergence and energy tests were used to give confidence in the results. The behavior of intensity and power near the tip were investigated. The effects of changing the dielectric constant of a sample material located below the tip were also examined.

1 Introduction

The *near-field*, in contrast to the *far-field*, refers to those spatial regions less than a wavelength from an aperture, obstruction, or surface which scatters electromagnetic waves. If harmonic time dependence is assumed, the Helmholtz equation governs spatial propagation in both regimes. In the far-field, approximation techniques such as Kirchhoff or Fresnel diffraction theory have been developed that allow the electric and magnetic fields to be computed with high accuracy. An understanding of the near-field, however, often requires a numerical solution to the Helmholtz equation. This is particularly true when the length scales of a problem extend from the near-field to the far-field and when geometries are complex. As the characteristic dimensions of features of technological significance have decreased, an understanding of near-field optical phenomena has become critically important. A particular example of this is provided by the tips used in Near-Field Scanning Optical Microscopes (NSOM).

Optical techniques have proved extremely useful in characterizing advanced materials and electronic devices. Diffraction, however, limits the spatial resolution that can be achieved with far field optics to dimensions comparable to the illumination wavelength. NSOM permits optical measure-

ments with much higher spatial resolution [1] and [2]. In the most common implementation, a sub-wavelength aperture is scanned across a sample while maintaining a separation also much less than the wavelength ($\sim$ 10nm) between sample and aperture. Using this approach, spatial resolutions comparable to the aperture size and much less than the wavelength can be achieved. The critical element in this form of NSOM is the aperture, which is typically prepared by tapering an optical fiber to a point [3]. Because the taper causes the diameter of the fiber core to shrink below the cutoff diameter for the propagating mode, light begins to escape from the core before reaching the tip. This enlarges the effective size of the light source and limits resolution. Most tip preparation processes reduce this effect by coating the outside of the tip with metal while leaving an uncoated, sub-wavelength ($\sim$ 50nm) aperture at the point. During a scan, topographic information and optical signals such as transmission, reflection, photoluminescence and Raman scattering are simultaneously recorded with spatial resolutions of 50nm or less. While the technique is proving to be quite useful, interpreting the results of NSOM measurements is not nearly as straightforward as are those for the far-field. For example, both propagating and evanescent fields emanate from the tip into the sample and the achievable resolution is critically dependent on how

these fields spread and couple. Scattering of evanescent modes into the far field is a topic of discussion [4] and [5] as is optical contrast, which arises solely from coupling to surface topography rather than to intrinsic changes in optical properties [6]. There is a clear need for a better understanding of the fields emanating from the tip to assist in interpreting measurements and optimizing microscope performance. In this paper, we discuss a Green's function approach to modeling the electric and magnetic fields from an NSOM tip.

Several groups have applied other approaches to this problem. For example, the Bethe-Bouwkamp model [7] and [8] of light emission from a circular aperture illuminated by a plane wave has been used to simulate an NSOM tip aperture [9] and [10]. Martin et al. have discussed a generalized field propagator approach for three dimensional modeling of scattered light in the near field [11]. Valee et al. have applied a similar approach to simulations of the contrast mechanism in NSOM when the tip is used as an illumination source [12]. Novotny et al. have used a multiple-multipole method to simulate the near-field region of a two-dimensional tip [13]. The results of our analysis are in good agreement with this latter work as discussed below. A longer summary and discussion of different methods used to model optical

properties of near-field tips can be found in [14].

Below we outline our Green's function formalism and apply it to a two-dimensional model of a near-field tip. Our approach uses Green's theorem to reduce the problem to surface integral equations, thereby lowering computational complexity by one-dimension. These surface integral equations are solved numerically. Because of the multilayered form of the geometries studied, the matrix formed by this method is quite sparse. This sparsity may ultimately allow more efficient numerical algorithms to be developed. A number of accuracy tests are made, including convergence tests and, to our knowledge, the first reported energy check on a near-field tip model. While energy conservation is a necessary although not sufficient constraint on the performance of any simulation of this type, it can be difficult to obtain the required accuracy in modeling near field tips, where the field emitted by the tip is several orders of magnitude below the incident signal. Proper termination of the integral equations was found to be essential to the energy check. The field emitted from a 50nm aperture is then modeled, and coupling of the near-field light into dielectric media with varying index of refraction is explored.

2 Theoretical Method

The first step in the construction of the solution for the electric field for a given problem is a determination of a model geometry. The results presented in this paper were computed using the model geometry in Figure 1. Although an NSOM tip is three-dimensional, we have chosen a two-dimensional model for the initial development of the method. This permits a scalar wave solution for TE (s) polarized electric fields in the different regions, which simplifies the development. Also, we are able to make comparisons with previous results for similar 2-dimensional problems [13]. In our model, we assumed a normally incident plane wave with a vacuum wavelength of 488 nm present in Region 1.

The specific geometry used in this model was chosen to match the two-dimensional (or cross sectional) structure of an NSOM tip. However, our model is capable of calculating the s polarized electric field for any two-dimensional geometry with arbitrary regions and boundaries with two constraints. First, the boundaries must be comprised of straight line segments. Second, the permittivity, ϵ, and permeability, μ, must be constant in a given region.

With specified model geometry and assuming harmonic time dependence

in the fields, the Helmholtz equation,

$$(\nabla^2 + k_i^2)u_i(\vec{x}) = 0, \tag{1}$$

is the spatial wave equation for the scalar field, $u_i(\vec{x})$, in the ith region with wave number, k_i, in that region. u_i can be interpreted as a scalar velocity potential. For s polarization, this means that $u_i = E_i/\mu_i$, where E_i represents the electric field.

For the Helmholtz equation in two-dimensions, the Green's function, $G_i(\vec{x}, \vec{x}')$, in the ith region for field point, $\vec{x}$, and source point, $\vec{x}'$, is[15]

$$G_i(\vec{x}, \vec{x}') = \frac{i}{4} H_0^{(1)}(k_i|\vec{x} - \vec{x}'|), \tag{2}$$

where $H_0^{(1)}$ is the zeroth order Hankel function of the first kind. Green's theorem may be applied to the model geometry in closed regions to obtain the integral relation

$$u_i(\vec{x}) = \oint_{s_i} ds' \left\{ u_i(\vec{x}') \left(\hat{n} \cdot \vec{\nabla}' G_i \right) - G_i \left(\hat{n} \cdot \vec{\nabla}' u_i(\vec{x}') \right) \right\}, \tag{3}$$

where the integral is over the closed curve, s_i, that bounds the ith region, and $\hat{n}$ is the unit vector normal to the curve into the region. In our analysis, the term "boundary" designates a portion of a curve that is of constant slope. As can be observed from our geometry, the curves bounding the different regions frequently possess several "boundaries."

Eq. 3 requires that the field on the boundaries of a region be determined before the field for points inside the region can be calculated. To obtain an analytic expression for the field on the boundaries, we take the limit of $u_i(\vec{x})$ as $\vec{x}$ approaches the boundary of Eq. 3. The boundaries will be approached using the limit from both the region above and the region below. In so doing, the surface integrals contain integrable singularities (when the field and source points coincide); the field and its normal derivative on the boundaries are not singular. The part of the integrals involving the Green's function have logarithmic singularities that integrate to zero. However, the parts of the integrals involving the normal derivative of the Green's function contain singularities that do contribute. This is handled by integrating a small arc around the singular point, and taking the limit as the radius of the arc goes to zero. In general, the result for this part of the integral is $u_i \Delta\theta/2\pi$, where $\Delta\theta$ is the angular distance traversed during the small arc integration. The straight line components of the surface integrals will yield $\Delta\theta = \pi$. For corners, $\Delta\theta$ will be the angle subtended by the two boundaries outside the region of interest. See [16] for more details on this limiting procedure.

The use of Green's theorem requires a closed surface integral bounding a region to yield the field in that region. As can be seen in our geometry, none

of the regions are closed. Some consistent method must be used to close the different regions.

For Region 1, the integral over the curves bounding the region is split into four different pieces. The first piece is the integral over the boundaries in Figure 1 with $-R \leq x \leq R$. The second piece involves extending the boundary between Region 1 and Region 3 from $+R$ to $+\infty$. We make the assumption that the integration is on a curve far enough away from the tip that the field has settled down to that of a normally incident plane wave striking an infinite, flat boundary. This field (and its normal derivative) can then be determined using the Fresnel equations. The integral for this segment in Eq. 3 then takes the form

$$I_i^+(x,y) = \int_R^\infty dx' \left\{ u_0(\vec{x}') \left(\hat{n} \cdot \vec{\nabla}' G_i \right) - G_i \left(\hat{n} \cdot \vec{\nabla}' u_0(\vec{x}') \right) \right\}, \qquad (4)$$

where u_0 is the total (reflected + incident) field given by the Fresnel equations, and $i = 1$ for Region 1. This integral is calculated numerically. The third piece of the integral involves extending the boundary between Region 1 and Region 2 from $-R$ to $-\infty$. We make the same assumptions on the field and its normal derivative as in the previous piece. By symmetry, this integral is similar to the integral from $+R$ to $+\infty$. The differences are that the limits are now from $-\infty$ to $-R$, and the x coordinate is replaced by

$-x$. This piece of the integral will subsequently be referred to as $I_i^-(-x, y)$. The assumed termination of the field for the surface integrals for $x > R$ and $x < -R$ was tested and found to be an excellent approximation as discussed in the Confidence Tests section of the paper. The final piece of the integral is a half-circle that extends from $+\infty$ in x to $+\infty$ in y to $-\infty$ in x. This half circle integration will evaluate to the incident field. For the details of this calculation for infinite curves, see [17] and [18].

For Region 2, the upper boundary has already been extended to $-\infty$, yielding $I_i^-(-x, y)$ as previously discussed. The only difference is that $i = 2$ in Eq. 4. The lower boundary (separating Region 2 from Region 4) is also extended from $-R$ to $-\infty$. In this case, we assume that the field is zero, and thus the integral from $-\infty$ to $-R$ does not contribute. This assumption is justified by examining the results of the calculation. The calculated field on this curve near $x = R$ was less than 10^{-10} of the incident field. Region 2 is then closed with a vertical segment that joins the two boundaries at $-\infty$. This segment also does not contribute. Region 3 is closed in a manner similar to Region 2 but at $+\infty$.

Region 4 is terminated in a manner similar to Regions 2 and 3. The boundaries are extended to $\pm\infty$ with the assumption that the field for $x > R$

and $x < -R$ is zero. This causes the integrals for $x > R$ and $x < -R$ to be zero. Integrals over vertical segments at $\pm\infty$, which do not contribute, close the region.

In Region 5, the horizontal boundary has already been extended to $+\infty$ and $-\infty$ as described above. The region is closed by a half-circle that extends from $-\infty$ in x, to $-\infty$ in y, to ∞ in x. Because the transmitted field in Region 5 satisfies a Sommerfeld radiation condition, this half-circle integral will not contribute.

The net result for the field along the boundaries will be the integral equation

$$\frac{\Delta\theta}{2\pi}u_i(\vec{x}) = u_{inc}(\vec{x}) + I_i^+(x,y) + I_i^-(-x,y) + \int_{s_i} ds' \left\{ u_i \left(\hat{n} \cdot \vec{\nabla}' G_i \right) - G_i \left(\hat{n} \cdot \vec{\nabla}' \right) u_i \right\}. \tag{5}$$

The remaining integral in this equation will now only involve the boundaries that are actually in the model geometry ($-R \leq x \leq R$). In the integral, u_i will be a function of $\vec{x}'$, and G_i will be given by Eq. 2. Also, $u_{inc}(\vec{x})$ is only present for Region 1 boundaries. $I_i^+(x,y)$ will be present only for Region 1 and 3 boundaries. While $I_i^-(-x,y)$ occurs only for Region 1 and 2 boundaries.

The integral of Eq. 5 must be evaluated numerically through some dis-

cretization technique. We chose n discretization points, which break the curve into n-1 segments. The parameter n must be sufficiently large that the segments are much smaller than the incident wavelength. By allowing a discretized point on the curve to act as a field point, while the other discretized points act as source points, an integral equation relating the field at the field point to the field and its normal derivative at the source points is generated from Eq. 5.

The integral in Eq. 5 evaluated in the regions above and below a boundary point produce 4 unknowns at each boundary point: field above, field below, normal derivative above, and normal derivative below. Discretization of the integral in Eq. 5 gives algebraic expressions involving the field and normal derivative unknowns. In our discretization approach, each term in the integrand is approximated on segments between boundary points as a polynomial (i.e. a power series). The discretization converts Eq. 5 to:

$$\frac{\Delta\theta}{2\pi}u_p = u_{inc}(x_p, y_p) + I_i^+(x_p, y_p) + I_i^-(-x_p, y_p) + \sum_{q=1}^{\mathrm{n}} \Delta l_q (a_{pq}u_q - b_{pq}N_q). \quad (6)$$

In this equation, the subscripts have been changed to represent the point at which the function is evaluated. All functions are evaluated in the region of the field point. u_p is now the field value at the pth boundary point (field point), (x_p, y_p); u_q is the field and N_q is its normal derivative at the qth

boundary point (source point). Below, u and N are referred to as boundary unknowns. Δl_q is the distance between successive boundary points (discretization length); a_{pq} is a term which is derived from the normal derivatives of the Green's function with field point p and source points near point q; b_{pq} is similarly obtained from the Green's function evaluated with field point p and source points around point q. The form of the latter two variables depends on the polynomial approximation. For example, in the simplest (zeroth order polynomial) approximation, the boundary unknowns, Green's function and normal derivatives of the Green's function, are assumed constant between boundary points. In this case, a_{pq} and b_{pq} equal the Green's function and the normal derivative of the Green's function with field point p and source point q, respectively.

When more terms are added to the power series, a_{pq} and b_{pq} are no longer simply the values of the Green's function and its normal derivative. They are derived from linear combinations of the Green's function and its normal derivative evaluated with source points at the qth boundary point and neighboring boundary points (all using the same field point). For example, a first order (linear) power series expansion in the distance, l, from the center of a

segment would yield

$$f(l) = \frac{1}{2}(f_q + f_{q+1}) + \frac{1}{\Delta l_q}(f_{q+1} - f_q)l, \tag{7}$$

where f represents either a boundary unknown, the Green's function, or its normal derivative. The dependence of the Green's function and its normal derivative on the field point are constant over the integration, so these dependences have been omitted. Therefore, given a field point, f_{q+1} is the function value at the $(q+1)$th source point and f_q is the function value at the qth source point. To approximate the integral in Eq. 5, the power series for the field and the normal derivative of the Green's function are multiplied and analytically integrated on the variable l over the length of the segment (Δl_q). The terms multiplying the field unknown, u_q, are collected to obtain a_{pq}. a_{pq} then becomes a linear combination of the normal derivative of the Green's function evaluated at the qth point and surrounding points which were used to define the power series. The multiplication and integration is repeated for the normal derivative of the field and the Green's function to obtain b_{pq}. This multiplication and integration of power series is repeated for all of the n-1 power series definitions on the curve to obtain a general expression for all the a_{pq}'s and b_{pq}'s. Convergence of the boundary unknowns was greatly improved by increasing the order of the polynomial approximation.

All of the results presented here used quadratic approximations for each of the functions as long as the argument of the Green's functions and its normal derivative are sufficiently large.

The singularities at zero in the Green's function and its normal derivative necessitate a different power series to calculate the integrals when the argument is small (we chose the cutoff to be $\sim 0.7|2\pi/k_i|$). As a small argument expansion, we used the definition of the Hankel functions given by [19]. This form of the functions consists of combinations of power series and logarithms. These are multiplied by the power series for the boundary unknowns used above and, again, integrated to obtain values for the a_{pq}'s and b_{pq}'s. This method is efficient due the fast convergence of the power series for the Hankel functions for small argument.

If this procedure is repeated n times, allowing all of the discretized points on a boundary to act as field points, 2n equations and 4n boundary unknowns result. The process is repeated for each of the boundaries in the geometry. Therefore, if m total boundary points are chosen, a linear system of 2m equations and 4m boundary unknowns results. The linear system requires 2m additional constraint equations. To determine the additional constraints, continuity conditions are enforced. Every boundary separates two regions, i

and j. For s polarized light, the boundary conditions require that

$$\begin{aligned} \mu_i u_i &= \mu_j u_j \\ N_i &= N_j. \end{aligned} \tag{8}$$

These relations provide the $2m$ additional constraint equations.

For the integration involving the region below the field point, the continuity equations are used to eliminate the field and normal derivative of the field below in favor of the field and normal derivative of the field above. This will introduce a factor of the relative permittivity between the two regions. In this case, the relative permittivity factor will be absorbed into the a_{pq}'s and b_{pq}'s.

Now the numerical approximation to the set of integral equations can be written as a well defined linear system involving the boundary unknowns. All known quantities in Eq. 6 are separated from the terms containing the boundary unknowns and the linear system is expressed in an impedance matrix formalism

$$Z\vec{b} = \vec{u}_{inc} + \vec{I}^{+} + \vec{I}^{-}. \tag{9}$$

Here, $\vec{b}$ is a vector whose first half contains the unknown fields (u_q) and the second half contains the unknown normal derivatives (N_q) at the discretized points on the boundaries. The elements of $\vec{u}_{inc}$ are the incident field value at

the discretized points in Region 1 and zero otherwise. The elements of the vector $\vec{I}^{+}$ will be the values of the integral calculated from Eq. 4 for points on the boundaries in Regions 1 and 3, and zero for points in the other regions. Similarly, the elements of $\vec{I}^{-}$ will be the values of the integral calculated from Eq. 4 for points on boundaries in Regions 1 and 2, and zero otherwise. Z is a generalized impedance matrix.

For the impedance matrix, each row dotted into $\vec{b}$, represents the discretized form of one of the 2m integral equations (Eq. 5), with one of the boundary points acting as the field point, $\vec{x}$. This causes Z to have 2m rows. Each boundary point will correspond to two rows, i.e. it will act as a field point twice (bounds two regions). The upper half of the impedance matrix represents Eq. 6 for every field point with the integration in the region above. The lower half repeats the integrals for each field point using the region below in the integrals. For each column, the boundary points are acting as the source points, $\vec{x}'$, in the integrals of Eq. 5. The matrix elements for the pth field point and qth source point are therefore given by

$$z_{pq} = \begin{cases} a_{pq}\Delta l_q - \frac{\Delta\theta}{2\pi}\delta_{pq} - \frac{\mu_i}{\mu_j}\frac{\Delta\theta}{2\pi}\delta_{(p+\mathrm{m},q)}; & q \leq \mathrm{m} \\ b_{pq}\Delta l_q; & q > \mathrm{m} \end{cases} \tag{10}$$

where δ is the Kronecker delta function, and i and j are the region numbers

for the region above and below the boundary, respectively, for field point p. a_{pq} and b_{pq} are zero when the boundary on which the source point lies does not bound the region for row p. This leads to a relatively sparse matrix (roughly half of the elements are zero for this geometry). In our present implementation, we have not focused on developing algorithms to take advantage of this sparsity. In a 3-dimensional implementation, however, developing such algorithms could lead to significant improvements in computational speed.

After the elements of Z have been computed, the linear system is solved using an LU (Lower/Upper Triangular) decomposition to obtain the boundary unknowns. The boundary unknowns are then substituted into Eq. 3. This equation is then numerically integrated yielding the field at any point in a given region.

Once the field has been obtained, an external method that can check the consistency of the calculated fields is used to provide an independent check of the the accuracy of the fields. We have devised an energy check to give an indication of energy balance in the system. We calculate the net flux into a closed box and the energy dissipated in the absorbing regions inside the box. Energy conservation constrains these two values to be equal. The difference between these two values provides an estimate of the error in the calculated

fields.

For the harmonic fields, u_i, the Poynting vector takes the form

$$\vec{S} = \frac{i\mu_i^*}{2\omega}\left(u_i\vec{\nabla}u_i^*\right). \tag{11}$$

The (integrated) flux through a line (l) is then given by

$$\int_l \mathrm{Re}\left(\vec{S}\right)\cdot\hat{n}_l dl, \tag{12}$$

where $\hat{n}_l$ is the unit normal to the line. This equation is used to calculate the net flux through the four sides of a box. The dissipated (absorbed) energy is obtained by calculating the net flux through the boundaries of the absorbing regions within the box using Eq. 12. Another calculation of interest is of integrated intensity. The integrated intensity is Eq. 12 with $\mathrm{Re}\left(\vec{S}\right)\cdot\hat{n}_l$ replaced by uu^*. This calculation is presented in the Results section of the paper as an indication of the presence of evanescent modes.

In addition to lending confidence to the accuracy of the calculated fields as discussed below, the energy check provides a method for examining energy flow in the system. It can therefore provide an indication of the efficiency of a given tip configuration. This information can assist in the optimization of the properties of tips. The energy check can also yield valuable information concerning the behavior of energy transport in different dielectric media.

This is an important step in the characterization of the properties of different materials.

Although we emphasize s polarized radiation in this paper, the method has a natural extension for the solution of the TM (p) polarized problem which we outline here for completeness. The analysis follows a relatively similar thread, with only a few changes. For p polarization, the direction of $\vec{H}$ is conserved. For this reason, the interpretation of u_i would change to $u_i = H_i/\epsilon_i$. Maxwell's equations relate the electric field to u_i by

$$\vec{E}_i = \frac{i}{\omega}\left(\frac{\partial u_i}{\partial y}\hat{x} - \frac{\partial u_i}{\partial x}\hat{y}\right), \tag{13}$$

where ω is the angular frequency of the incident wave.

The formulation of the integral equations and the Green's function for the p polarization would be identical to that for s polarization. The continuity conditions would take on the form

$$\begin{aligned} \epsilon_i u_i &= \epsilon_j u_j \\ N_i &= N_j. \end{aligned} \tag{14}$$

The Poynting vector also takes on a slightly modified form

$$\vec{S} = \frac{i\epsilon_i^*}{2\omega}\left(u_i \vec{\nabla} u_i^*\right). \tag{15}$$

For the p polarized field, the closure technique for Region 1 is somewhat

more problematic. As one moves away from the tip, there will be a component of the electric field that is perpendicular to the curve. These waves do not excite surface currents, and therefore propagate asymptotically as $1/\sqrt{r}$. This decay is slow enough that R would have to be taken to a greater distance than in the s polarization case, or an alternative termination scheme would need to be developed.

3 Results

3.1 Confidence Tests

We directly tested the accuracy of the calculational method in several different forms. First, we analyzed the correctness of the use of the asymptotic (infinite flat boundary) value for the integrals from R to ∞. We also examined the convergence behavior of the boundary unknowns with respect to increasing boundary point density. A check on the conservation of energy in the model was also performed.

We tested the accuracy of using asymptotic values of the boundary unknowns at the edges of the geometry in Region 1. By symmetry, the analysis at $x = +R$ will be identical for that of $x = -R$. For $R = 2000$ nm, the

boundary unknowns near R converged to within 1% of the asymptotic value given by the Fresnel equations. The consistency with the asymptotic value is most influential to the field local to R. However, we are most interested in the field near the tip of the NSOM. This warrants a check of the convergence of the boundary unknowns in the region near the tip as a function of R. The field and normal derivative showed a relative change of less than 2×10^{-4} as R changed from 2000 nm to 8000 nm. This indicates that the closure method yields a stable field near the tip that has essentially converged with a relatively modest value of 2000 nm for R.

We also examined convergence of the boundary unknowns on the tip as a function of boundary point number. We first attempted to optimize the rate of convergence by balancing the error in the integrals. Because of the damped nature of the fields near the tip, it is plausible that the fields are more ill-behaved there. For this reason (and the fact that the field near the tip is of paramount importance), we used higher point densities near the tip. To ensure symmetry about the y axis was preserved, we chose boundaries symmetrically located about the y axis to have the same boundary point density. Other than this symmetry, none of the boundaries possessed identical densities. In addition, several of the boundaries near the tip were

split into a segment close to the tip (of higher point density) and a segment farther from the tip (of lower point density). The initial point densities were roughly: every 5 nm for segments closest to tip, every 15 nm for segments of intermediate distance (250 nm to 1000 nm from tip), and every 30 nm for segments farthest from the tip. Once the initial densities for the boundaries were chosen, convergence was tested by increasing the boundary point densities on each segment by the same multiplicative factor. Table 1 shows the convergence of the field and the normal derivative of the field at the center of the tip aperture as a function of the total number of boundary points. In this table, B.P.# is the total number of boundary points, u is the field unknown at the center of the tip, N is the normal derivative unknown at the center of the tip, and $\Delta\%$ is the percent change in the unknown from the previous calculation.

We also tested the program's consistency with energy conservation. We calculated the net flux into a box surrounding the tip. The box was 1400 nm wide, 1400 nm tall, extended 200 nm below the end of the tip and was symmetrically placed in the horizontal direction. This box is labeled C1 in Figure 1. The excess energy entering the box not absorbed is 0.0014% of the incident flux through the top of the box. This is 0.70% of the power emitted

from the tip. This indicates that the fields below the tip are quite accurate with respect to energy conservation, increasing our overall confidence in the correctness of our calculations.

3.2 Calculated Results

To explore the energy emitted by the tip, we constructed three sides of a box that enclosed the tip, shown as C2 in Figure 1. For this box calculation, the sample was removed (no Region 5), leaving vacuum below the tip. The bottom of the box was then located progressively farther from the tip. The integrated intensity and the net flux for the two sides and the bottom was calculated. The results of this analysis are presented in Figure 2. The net flux through the sides and the bottom of the box remains constant as the bottom of the box is moved from the tip. In contrast, the integrated intensity decays as the bottom is moved away. This is a clear signature of the presence of both evanescent and propagating fields in the region below the tip: evanescent fields contribute to the intensity but do not propagate into the far field and hence do not contribute to the flux. The slight increase in integrated intensity observed at approximately 2000 nm is an artifact of the integration geometry. The integration is along a semi-box and not a semi-circle. The intensity

asymptotically approaches $1/r$ behavior for large distances from the tip, and the pathlength along a semi-circle integration behaves like r. The net result would be that the integrated intensity along a semi-circle would approach a constant. For a box, however, the integration path is now along a line. The $1/r$ behavior will lead to a divergent integral as the length of the line approaches ∞(i.e. as the length of the side integrals go to ∞).

A contour plot of $\log(|E|^2)$ is illustrated in Figure 3, where the sample has been put back in place. We chose the sample to have an index of 1.5 to allow for comparison of our results with those of Novotny et al. [13]. The results we obtained are, in form, the same as those obtained by Novotny. This figure also contains an inset that displays a vertical cut showing the decay of the intensity (linear scale) from the tip. Note that the intensity drops more rapidly with distance than the integrated intensity in Figure 2 because the former includes both the decrease from the evanescent field and from the 1/r geometrical factor.

We altered the index of refraction in the lower medium, and investigated the results to learn more about how achievable resolution depends on the index of the sample being studied. For this analysis, three different indices of refraction were used, corresponding to vacuum (essentially no medium

present below the material), glass, and a typical semiconductor, GaAs. Figure 4 displays a contour plot of $\log(|E|^2)$ for both the near-field and far-field radiation patterns on the same scale for the different cases we examined. In each panel, the contour lines differ by a factor of $10^{0.6}$. The lowest contour in each of the plots is $10^{-8.6}$. The lower aperture was 50 nm, while the upper aperture was 850 nm. Case (a) corresponds to vacuum below the tip. In this case, the radiation spreads significantly in all directions. In (b), glass (index of 1.5) was placed 50 nm below the lower tip aperture. While 50 nm is somewhat larger than typical tip-sample separation in NSOM, this value was chosen to allow the effects on the fields between the tip and sample to be more easily observed. The dashed line in (b) indicates the critical angle in the glass. As can be observed from the figure, the index change at the lower interface causes both the near and far fields to become much more confined. The effect is more evident in (c), which contains GaAs (index of 3.4). Again the dashed line represents the critical angle. In this plot, the fields are greatly confined, even relative to the glass plot. In situations where weak absorption is being detected, spreading of the field places a limit on spatial resolution. It is clear from Figure 4, that the higher index materials will have better spatial resolution.

Knowledge of the behavior of the transmission of power into the sample in the NSOM is of interest. Specifically, the amount of power transmitted as a function of the dielectric constant of the sample can yield important information regarding the efficiency of a given tip configuration for viewing a sample and also allow calibration of optical results from two samples with different indices measured with the same tip. Figure 5 displays the power transmitted into different media as a function of index. If there were no gap between the tip and the sample, and they both possessed the same index, no reflections would occur. It is reasonable to assume that adding a small vacuum gap will not significantly alter the transmission behavior as long as the gap is much smaller than the wavelength. The index corresponding to maximum transmission corresponds to a sample with an index close to that of the tip material. As can be observed from the figure, this is indeed the case.

4 Conclusions

This paper presents a novel approach for numerically modeling two-dimensional, s polarized, wave propagation problems that involve several different length

scales with minimal constraints. The method was applied to the problem of the Near Field Scanning Optical Microscope. Several convergence and energy tests were used to increase confidence that the results presented were accurate. Intensity and energy behavior near the tip were investigated. The effects of changing the dielectric constant of the sample material were also examined.

The results of this paper indicate that our Green's function approach can be effective in modeling the behavior of the electric field in near-field problems. The results give support for the continued development of the method, and possibly the extension to the more complex three-dimensional problems.

5 Acknowledgements

The authors gratefully acknowledge valuable discussions with John Flory, Mary Herndon, William Bradford, David Wood, Jim Bernard, and Paul Martin. This material is based upon work supported by the National Science Foundation under Grant No. DMR-9704780. Additional support was supplied from the Colorado Photonics and Optoelectronics Program and the

MRS UMRI program.

References

[1] U. Durig, D. W. Pohl, and F. Rohner, "Near-field optical-scanning microscopy," J. Appl. Phys. 59, 3318-3327 (1986).

[2] E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, "Breaking the diffraction barrier: Optical microscopy on a nanometric scale," Science 251, 1468-1470 (1991).

[3] G. A. Valaskovic, M. Holton, and G. H. Morrison, "Parameter control, characterization, and optimization in the fabrication of optical fiber near-field probes," Appl. Optics 34, 1215-1228 (1995).

[4] K. Fukuzawa, and H. Kuwano, "Conversion of evanescent into propagating light in near-field scanning optical microscopy," J. Appl. Phys. 79, 8174-8178 (1996).

[5] E. Wolf and J. T. Foley, "Do evanescent waves contribute to the far field?", Opt. Lett. 23, 16-18 (1998).

[6] L. Novotny, B. Hecht, and D. W. Pohl, "Implications of high resolution to near-field optical microscopy," Ultramicroscopy 71, 341-343 (1998).

[7] H. A. Bethe, "Theory of Diffraction by Small Holes" Phys. Rev. 66, 163-182 (1944).

[8] C. J. Bouwkamp, "On Bethe's theory of diffraction by small holes," Philips Res. Rep. 5, 321-332 (1950).

[9] G. W. Bryant, "Probing quantum nanostructures with near-field optical microscopy and vice versa," Appl. Phys. Lett. 72, 768-770 (1998).

[10] G. W. Bryant, E. L. Shirley, L. S. Goldner, E. B. McDaniel, J. W. P. Hsu, and R. J. Tonucci, "Theory of probing photonic crystal with transmission near-field optical microscopy," Phys. Rev. B15 58, 2131-2141 (1998).

[11] O. J. F. Martin, C. Girard, and A. Dereux, "Generalized field propagator for electromagnetic scattering and light confinement," Phys. Rev. Lett. 74, 526-529 (1995).

[12] P. J. Valle, R. Carminati, J-J. Greffet, "Contrast mechanisms in illumination-mode SNOM," Ultramicroscopy 71, 39-48 (1998).

[13] L. Novotny, D.W. Pohl, P. Regali, "Light propagation through nanometer-sized structures: the two-dimensional-aperture scanning near-field optical microscope." J. Opt. Soc. Am. A 11 No. 6, 1768-1779 (1994).

[14] D. W. Pohl and D. Courjon, eds., Near Field Optics, NATO ASI Series E, Vol 242 (Kluwer, Dordrecht, The Netherlands, 1993)

[15] DeSanto, John A. Scalar Wave Theory: Green's Functions and Applications, Springer, Heidelberg, 1992.

[16] Colton, D. and Kress, R. Integral Equation Methods in Scattering Theory, Wiley, New York, 1983.

[17] DeSanto, John A. and Martin, P.A. "On angular-spectrum representations for scattering by infinite rough surfaces," J. Wave Mot. 24, 421-433 (1996).

[18] DeSanto, John A. and Martin, P.A. "On the derivation of boundary integral equations for scattering by an infinite one-dimensional rough surface," J. Acoust. Soc. Am. 102, 67-77 (1997).

[19] Handbook of Mathematical Functions. Milton Abramowitz and Irene A. Stegun. New York: Dover, 1972.

Table 1: Percent change (Δ) from previous value of the field (u) and its normal derivative (N) at center of lower tip aperture as a function of total number of boundary points (B.P.#).

B.P.#	u	$\Delta\%$	$N \times 10^{-3}$	$\Delta\%$
585	$-0.11038 + 0.37843i$	-	$-0.084947 + 9.9202i$	-
1169	$-0.11025 + 0.37871i$	0.0771	$-0.085528 + 9.9367i$	0.166
1755	$-0.11020 + 0.37864i$	0.0215	$-0.085860 + 9.9392i$	0.0263

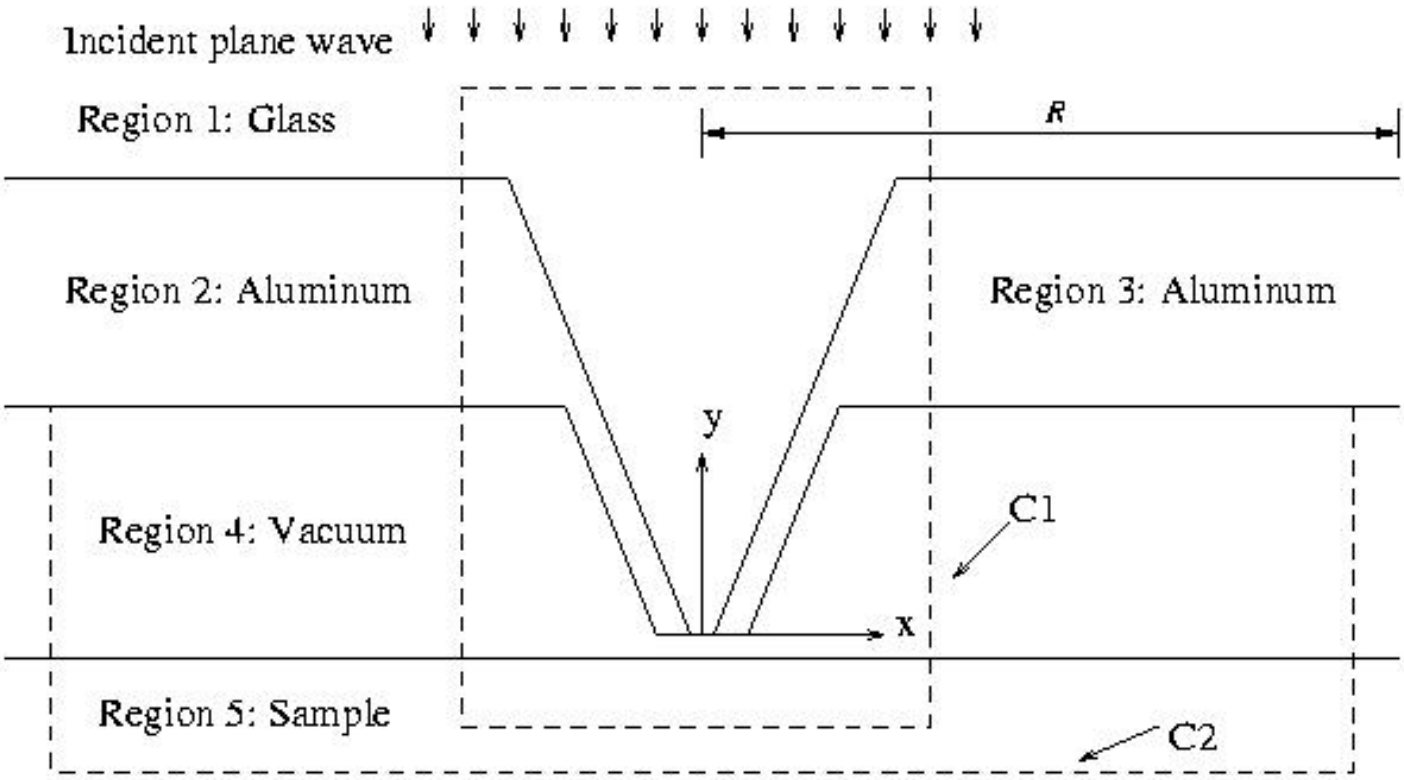
Figure 1: The geometry used throughout this paper to model the NSOM tip is shown. The upper aperture in region 1 is 850 nm. The lower aperture between regions 1 and 4 is 50 nm. The vertical distance between the upper aperture and the lower aperture is 1000 nm . The thinner aluminum strips along the taper of the tip have a thickness of 70 nm. The thicker aluminum strips have a thickness of 500 nm. The flux and integrated intensity are computed along C1 and C2 as discussed in the Results section of the paper.

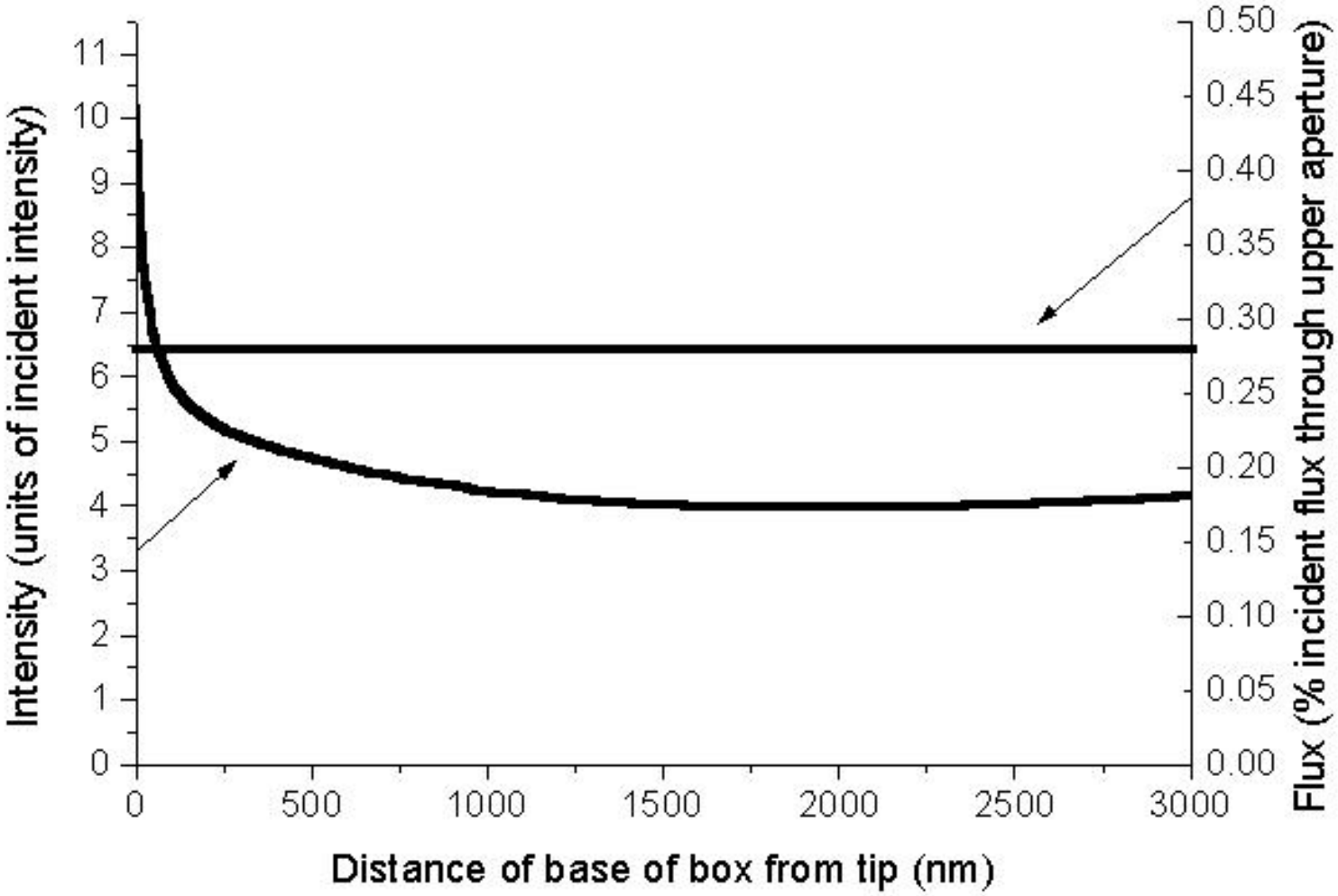
Figure 2: The integrated intensity along the curve C2 in Fig. 1 and the integrated flux (power) through C2 are shown for a geometry where the sample (Region 5) has been removed. The lower line of curve C2 was shifted progressively further from the tip. The distance from the tip is a measure of how far this lower line was from the lower tip aperture. The integrated intensity is normalized to the incident integrated intensity along the upper aperture. The power is normalized to the incident power through the upper aperture.

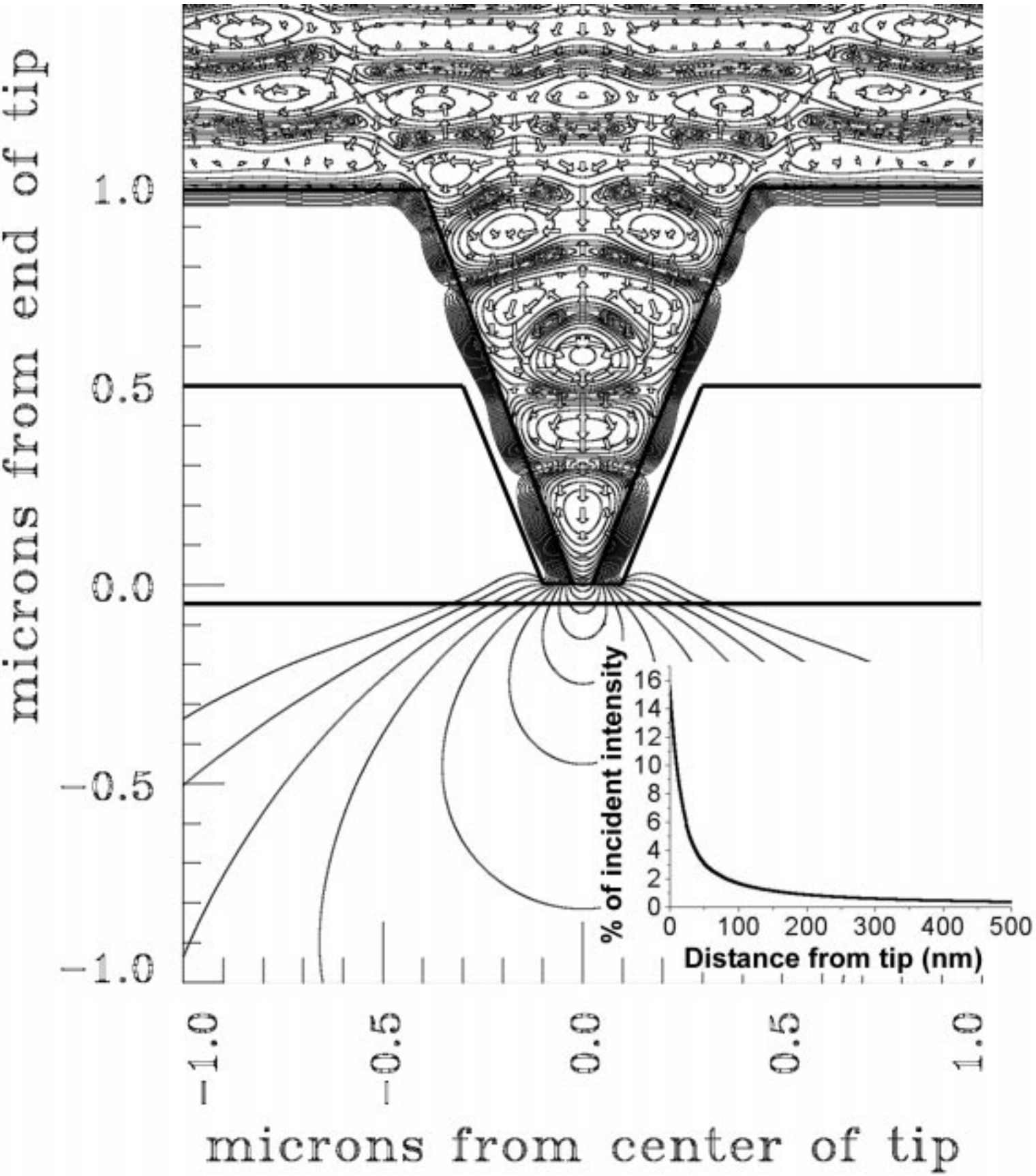
Figure 3: A contour plot of $\log(|E|^2)$ (logarithm of the electric field intensity) with glass (n=1.5) as the sample is shown. Poynting vectors have been included in the figure. Each contour line differs by the factor of $10^{0.6}$ with the lowest contour at $10^{-8.6}$ of the incident intensity. The inset is a linear plot of the decay of the percent intensity of the field (normalized to the incident intensity) along a vertical starting from the center of the lower tip aperture.

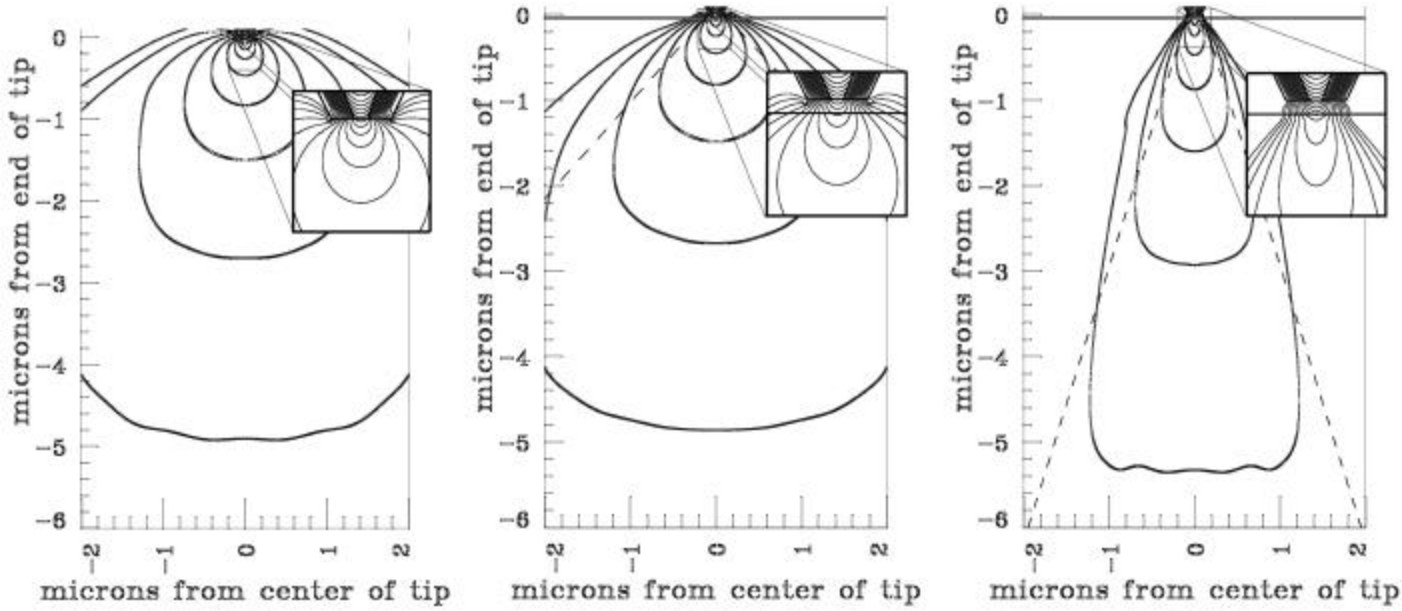
Figure 4: The far-field radiation intensity patterns on the same scale for three different cases is shown: (a) vacuum sample (n=1), (b) glass sample (n=1.5), (c) GaAs sample (n=3.4). The contours differ by a factor of $10^{0.6}$ where the lowest contour is $10^{-8.6}$ of the incident intensity. In (b) and (c) the dashed lines indicate the critical angles in the sample.

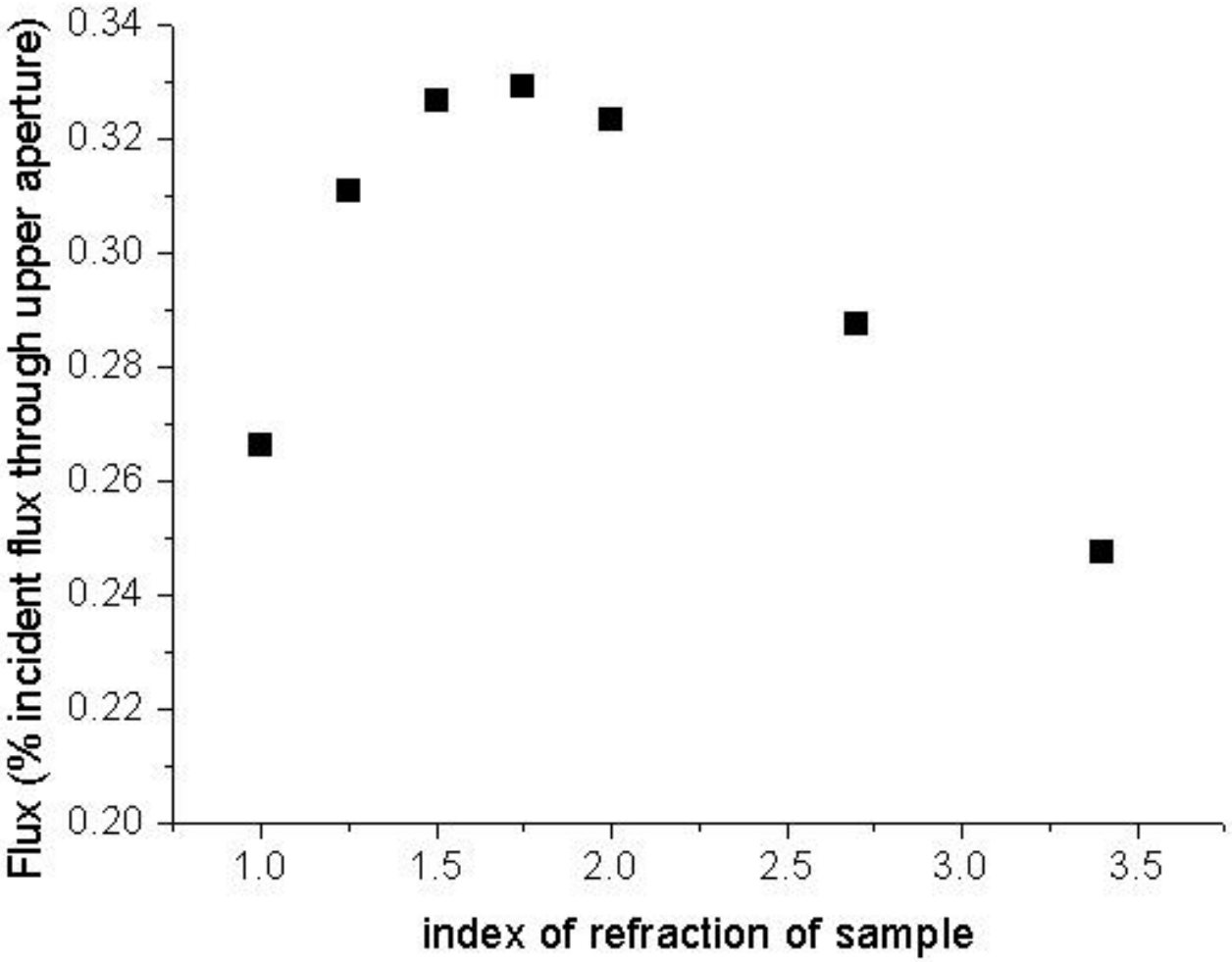
Figure 5: A plot of the integrated flux (power) into the sample as a function of the index of refraction of the sample material is shown. The power is normalized to the incident power through the upper aperture.


Incident plane wave
Region 1: Glass
R
Region 2: Aluminum
Region 3: Aluminum
y
Region 4: Vacuum
C1
x
Region 5: Sample
C2


Intensity (units of incident intensity)
Flux (% incident flux through upper aperture)
Distance of base of box from tip (nm)
0
1
2
3
4
5
6
7
8
9
10
11
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0
500
1000
1500
2000
2500
3000


microns from end of tip
1.0
0.5
0.0
−0.5
−1.0
1.0
0.5
0.0
0.5
1.0
microns from center of tip
% of incident intensity
16
14
12
10
8
6
4
2
0
0
100
200
300
400
500
Distance from tip (nm)


microns from end of tip
0
−1
−2
−3
−4
−5
−6
−2
−1
0
1
2
microns from center of tip
(a)
microns from end of tip
0
−1
−2
−3
−4
−5
−6
−2
−1
0
1
2
microns from center of tip
(b)
microns from end of tip
0
−1
−2
−3
−4
−5
−6
−2
−1
0
1
2
microns from center of tip
(c)


Flux (% incident flux through upper aperture)
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
1.0
1.5
2.0
2.5
3.0
3.5
index of refraction of sample