

Total Internal Reflection Tomography (TIRT) for Three-Dimensional Sub-Wavelength Imaging

David G. Fischer

NASA Glenn Research Center
MS 110-3
Cleveland, OH 44135

P. Scott Carney

Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61821

Foreword

Looking at the participants of the conference and the authors herein, one is struck by the fact that, beyond his specific scientific achievements and scholarly writings, Professor Wolf has amassed an amazing community of students, former students, collaborators and researchers who build upon the foundation that he has laid. This community shares an intellectual thirst, a dedication to scholarship, trust, respect and a somewhat unique sense of comradery. We would like to express our appreciation to Professor Wolf for our time as his students, our ongoing collaboration, and for introducing us to this community in its entirety. He has been and remains for us the prototype mentor, to be imitated but never truly copied in our own labs.

Abstract

In this chapter, we describe a novel form of near-field microscopy known as Total Internal Reflection Tomography (TIRT), which allows for three-dimensional sub-wavelength imaging. It is based on recent theoretical advances regarding the fundamental interaction of light with sub-wavelength structures, as well as stable algorithms for the

near-field inverse problem. We will discuss its theoretical underpinnings, as well as describe current efforts at NASA to implement a TIRT system for biofluids research.

1. Introduction

Near-field imaging has gained a great deal of exposure in recent years for its ability to resolve sub-wavelength structure in optically thin media [1-9]. It has many variants, including total internal reflection microscopy (TIRM) [5-7], photon scanning tunneling microscopy (PSTM) [9, 10], and near-field scanning optical microscopy (NSOM) [1-4], but common to all is the use of evanescent waves for illumination and/or detection. In many instances, image interpretation is difficult, owing to the complex interaction between the incident field and the sample, as well as between the scattered field and the near-field probe.

These difficulties are exacerbated when near-field techniques are applied to relatively thick samples. In addition to the problem of reconstructing a three-dimensional function of position (the dielectric susceptibility) from two-dimensional data sets (i.e. measurements of the scattered field in various planes), a thick object may exhibit strong scattering, with the consequence that the scattered field is a non-linear function of the susceptibility. Even when the scattering is weak, the detected field may not be simply related to the sub-wavelength structure of the object, as it is, for example, in the case of diffraction from a two-dimensional object [11, 12].

In this Chapter, we will discuss a new form of near-field imaging that makes use of TIRM measurements to produce computed reconstructions of the susceptibility of the sample. This method provides tomographic views and sub-wavelength resolution. Since

the system is free from the moving (and often ill-characterized) probe present in PSTM and NSOM, the analysis of the problem is greatly simplified. Indeed the experiment is well modeled as a half-space problem and an exact solution for the Green's function (absent the sample) is well known. The linearized inverse scattering problem may then be solved in a computationally efficient and stable manner. In Section II we review the fundamentals of diffraction tomography and observe the emergence of the classical resolution limits. In III we examine the properties of near-field evanescent waves and the role they play in achieving super-resolution in a variety of near-field methods. In IV we describe the basic TIRM measurement scheme and its extension to TIRT. In Section V we address the structure of the TIRT data and the development of fast, stable reconstruction algorithms followed by numerical simulations in VI. Finally, in Section VII we describe the instrument currently under construction at NASA to implement this modality.

2. Conventional Imaging

Conventional optical imaging systems are limited by diffraction to a resolution of approximately half the illuminating wavelength [13, 14]. This so-called Abbe-Rayleigh resolution limit [15, 16] is not fundamental, but is a consequence of the measurement scheme. In particular, the Abbe-Rayleigh limit arises only when the evanescent field in the near zone of the scatterer is inaccessible. For more than a century after the theoretical predictions of Abbe and Rayleigh, it was indeed the case that the so-called near field was not practically measurable.

Consider a scattering experiment in which a monochromatic field is incident on a localized dielectric medium with complex susceptibility $\eta(\mathbf{r})$ (see Fig. 1).

Figure 1. Illustrating the scattering geometry

For simplicity, we ignore the effects of polarization and consider the case of a scalar field $U(\mathbf{r})$ that obeys the reduced wave equation

$$\nabla^2 U(\mathbf{r}) + k^2 U(\mathbf{r}) = -4\pi k^2 \eta(\mathbf{r}) U(\mathbf{r}), \quad (1)$$

where k is the free space wave number. Following standard procedures, we find that $U(\mathbf{r})$ satisfies the integral equation [13]

$$U(\mathbf{r}) = U^{(i)}(\mathbf{r}) + k^2 \int d^3 r' G(\mathbf{r}, \mathbf{r}') U(\mathbf{r}') \eta(\mathbf{r}'), \quad (2)$$

where the outgoing Green's function $G(\mathbf{r}, \mathbf{r}')$ is given by

$$G(\mathbf{r}, \mathbf{r}') = \frac{\exp(i k |\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}, \quad (3)$$

and $U^{(i)}(\mathbf{r})$ is the incident field. We restrict ourselves to the weak-scattering approximation (also known as the first Born approximation [13]), which is particularly suited to the study of sub-wavelength structures. Accordingly, the scattered field $U^{(s)}(\mathbf{r}) = U(\mathbf{r}) - U^{(i)}(\mathbf{r})$ may be calculated perturbatively to lowest order in η with the result

$$U^{(s)}(\mathbf{r}) = k^2 \int d^3 r' G(\mathbf{r}, \mathbf{r}') U^{(i)}(\mathbf{r}') \eta(\mathbf{r}'). \quad (4)$$

We take the incident field to be a unit amplitude plane wave traveling in the direction of the unit vector \mathbf{s}_0 , i.e. $U^{(i)}(\mathbf{r}) = \exp(i\mathbf{k}\mathbf{s}_0 \cdot \mathbf{r})$. Utilizing the asymptotic form of the outgoing Green's function given by

$$\frac{\exp(ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} \sim \frac{\exp(ikr)}{r} \exp(-i\mathbf{k}\mathbf{s} \cdot \mathbf{r}'), \quad (5)$$

as $kr \rightarrow \infty$ with the unit vector \mathbf{s} kept fixed, we find that the scattered field in the far zone has the form

$$U^{(s)}(\mathbf{r}\mathbf{s}) \sim \frac{\exp(ikr)}{r} a(\mathbf{s}, \mathbf{s}_0), \quad (6)$$

where the scattering amplitude is given by the expression [13]

$$a(\mathbf{s}, \mathbf{s}_0) = k^2 (2\pi)^3 \tilde{\eta}[\mathbf{k}(\mathbf{s} - \mathbf{s}_0)], \quad (7)$$

and

$$\tilde{\eta}(\mathbf{K}) = \frac{1}{(2\pi)^3} \int d^3\mathbf{r} \eta(\mathbf{r}) \exp(-i\mathbf{K} \cdot \mathbf{r}) \quad (8)$$

is the three-dimensional spatial Fourier transform of the dielectric susceptibility. This is the fundamental result of diffraction tomography and was first derived by Professor Wolf in 1969 [14].

Equation (7) illustrates that, for a weakly scattering medium, there is a one-to-one mapping between the scattering amplitude for real \mathbf{s}, \mathbf{s}_0 and the low spatial frequency components of the dielectric susceptibility. Specifically, for a fixed direction of incidence \mathbf{s}_0 , the scattering amplitude is mapped onto the surface of a sphere of radius k centered at $-\mathbf{k}\mathbf{s}_0$ in the three-dimensional Fourier space of the dielectric susceptibility (Fig. 2).

Figure 2. Illustrating the Ewald sphere of reflection and the Ewald-limiting sphere.

As \mathbf{s}_0 is varied, those surfaces fill a sphere of radius $2k$ centered at the origin (known as the Ewald limiting sphere [13, 17]). Consequently, one may obtain a low-pass filtered estimate of the susceptibility, namely

$$[\eta(\mathbf{r})]_{\text{LP}} = \int_{|\mathbf{K}| \leq 2k} d^3\mathbf{K} \tilde{\eta}(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{r}). \quad (9)$$

Furthermore, this estimate is unique since, due to the analyticity of $\tilde{\eta}(\mathbf{K})$,¹ the low spatial frequency components of the dielectric susceptibility can, in principle, be analytically continued to the exterior of the Ewald limiting sphere [18]. In practice, however, techniques based on analytic continuation, such as band-limited extrapolation, are known to be unstable in the presence of measurement noise [19-21].

3. Evanescent Wave Illumination

We have seen that the low spatial frequency components of the susceptibility may be determined by illuminating a sample with homogeneous waves and measuring the scattered far field. To determine some subset of the Fourier components that lie outside

¹ It follows from the three-dimensional version of the Plancherel-Polya theorem that, since the domain of localization of the medium is finite, the three-dimensional spatial Fourier transform of the dielectric susceptibility is the boundary value on the real axes K_x, K_y, K_z of an entire analytic function of three complex variables.

the Ewald limiting sphere and consequently improve resolution, the illuminating field (or the measured field) must contain non-negligible evanescent field components.

Consider a general plane-wave form for the incident field:

$$U^{(i)}(\mathbf{r}) = \exp[i\mathbf{k}(\mathbf{s}_{0\perp} \cdot \boldsymbol{\rho}) + s_{0z}z], \quad (10)$$

where $\mathbf{r} = (\boldsymbol{\rho}, z)$, $\mathbf{s}_0 = (\mathbf{s}_{0\perp}, s_{0z})$ and

$$s_{0z} = \begin{cases} \sqrt{1 - s_{0\perp}^2} & \text{when } |\mathbf{s}_{0\perp}| \leq 1 \\ i\sqrt{s_{0\perp}^2 - 1} & \text{when } |\mathbf{s}_{0\perp}| > 1. \end{cases} \quad (11)$$

For $|\mathbf{s}_{0\perp}| \leq 1$, the incident field is a homogeneous plane wave propagating in the direction of the unit vector \mathbf{s}_0 , as seen in the previous section. It has a transverse scale length

$$\lambda_{0\perp} = \frac{2\pi}{k_{0\perp}} = \frac{\lambda}{|\mathbf{s}_{0\perp}|} \quad (12)$$

that is larger than the wavelength (see Fig. 3).

Figure 3. Illustrating the transverse scale length of homogeneous and evanescent waves.

For $|\mathbf{s}_{0\perp}| > 1$, the incident field is an evanescent plane wave with a transverse scale length that is smaller than the wavelength (i.e. $\lambda_{0\perp} < \lambda$). Consequently, evanescent incident fields can probe and *encode* structure on spatial scales smaller than the illuminating

wavelength. Unfortunately, evanescent waves decay exponentially from their point of origin with decay rate

$$\gamma = \frac{2\pi}{\lambda} \sqrt{|\mathbf{s}_{0\perp}|^2 - 1}, \quad (13)$$

so that a two-fold increase in probe scale length ($|\mathbf{s}_{0\perp}| = 2$), for example, requires that the source of evanescent waves be located within a distance of $\lambda/10$ of the scattering structure. This is the reason the resolution of a conventional imaging system is $\lambda/2$; the source and detector are many wavelengths from the scattering medium, so that effectively, only homogeneous waves are present.

Evanescent waves for illumination may be generated by total internal reflection or by diffraction at a sub-wavelength aperture. The near-field scanning optical microscope (NSOM) is an instrument that utilizes the second mode of generation to locally confine a probe field at the surface of a sample (see Fig. 4) [1-4].

Figure 4. Illustrating the principle of NSOM. Z_s represents the fiber-sample separation.

Scanning the aperture across the sample and recording either the throughput or the reflected field as a function of probe position produces an image with sub-wavelength resolution. This technique was first proposed by Synge in 1928 [22] and experimentally realized using microwaves by Ash and Nicholls in 1972 [23]. Today, a tapered optical fiber with sub-wavelength tip cross-section w is typically used in place of the aperture for illumination (Fig. 4). When the fiber tip is very close to the surface of the sample, the

resolution is on the order of the tip size. As the fiber is removed from the sample, the localization of the incident light (and the corresponding resolution) is reduced due to the loss (decay) of the evanescent waves. Figure 5 illustrates some of the first experimental images taken with an NSOM [1]. One can clearly see the loss of resolution as the fiber tip is incrementally distanced from the sample.

Figure 5. AT&T NSOM images (from Ref. [1]). The fiber-sample separations are (a) near contact, (b) 5 nm, (c) 10 nm, (d) 25 nm, (e) 100 nm, and (f) 400 nm.

NSOM images have striking sub-wavelength detail, but their interpretation is often problematic, especially for thicker samples. This is due to the fact that an NSOM image is a measure of the field *outside* the sample, and the fundamental relationship (i.e. mapping) between the field and the optical properties of the medium is usually not taken into account. Rather, the mapping is modeled effectively as a (one-view) projection or shadowgram. Furthermore, an NSOM image is typically defocused since the probe field contains a continuous distribution of both homogeneous *and* evanescent waves and, consequently, its exact composition (and the resulting image) is a function of the fiber-sample separation.

This is illustrated in Fig. 6. We consider a three-dimensional object composed of two stacked planes, one at $z = 0.005\lambda$ and another at $z = 0.405\lambda$ ($z = 0$ being the location of the evanescent wave source). Figures 6a and 6b illustrate the object structure

in the two planes, respectively, and Figs. 6c and 6d illustrate their individual far zone images.

Figure 6. Illustrating NSOM imaging of a thick object.

These are qualitatively similar to the images in Figs. 5a and 5f obtained from the AT&T NSOM. It is clear that the object structure contained in the plane $z = 0.405\lambda$ cannot be resolved from far zone scattered field measurements. Furthermore, for the composite object, the image (i.e. scattered field) of the unresolved structure in the plane $z = 0.405\lambda$ will obscure and degrade the image of the structure in the plane $z = 0.005\lambda$.

It is clear from Fig. 6 that the field scattered from the deeper plane is simply defocused. This phenomenon is not particular to the near field, but since a physical lens for the near field is not currently available (due to the irretrievable loss of evanescent waves), a post-processing solution must be sought. That this need seemed to be so long overlooked might be explained by the fact that the data obtained at closest approach so much resembles an actual image. As we shall see in the next section, there exist modalities for which processing of the collected data is absolutely required.

4. Three-Dimensional Near-Field Imaging

Total internal reflection microscopy (TIRM) also involves illuminating a sample by evanescent waves [5-7]. In TIRM, an evanescent wave generated by total internal reflection illuminates an object and the scattered (or radiated) light is collected by a

standard microscope objective. Due to the exponential decay of the incident evanescent wave, the interrogation (or excitation) volume is limited in depth to a thickness of $z_p = 1/\gamma$, with more evanescent fields yielding narrower excitation regions [see Eq. (13)]. Consequently, TIRM provides a far-field (i.e. diffraction limited) intensity image of a sub-wavelength region near the exit face of the prism. While it does not provide sub-wavelength imaging, it does provide sub-wavelength localization in depth. As such it is extremely useful for surface inspection. TIRM can be taken a step further if images are recorded for a series of distinct evanescent incident fields. In this case, each image is the result of a different (and unique) exponential weighting of the susceptibility, and the composite image stack can be inverted (possibly) to provide sub-wavelength resolution in depth (i.e. the z -dimension). In either the standard or extended TIRM case, the *transverse* resolution is limited to $\lambda/2$ since the excitation volume is not constrained in the transverse dimension and only the amplitude of the scattered field is measured in the far zone.

Total internal reflection tomography (TIRT) [24-27] is a coherent extension of TIRM in which the complex scattered field (both amplitude and phase) in the far zone is recorded for a diverse set of evanescent incident fields. The scattered field data is subsequently synthesized (or inverted) to yield an estimate of the three-dimensional object structure. Central to the inversion is the physical connection between the structure of the susceptibility and the scattered field. This connection is most easily seen in the Fourier domain. Consider an evanescent plane wave with complex wave-vector $\mathbf{k}\mathbf{s}_0$, generated by total internal reflection, which illuminates a weakly scattering object that is confined to the region $0 \leq z \leq d$ (see Fig. 7).

Figure 7. Illustrating the TIRT geometry. θ_c represents the critical angle.

The amplitude of the evanescent wave at the exit face of the prism (i.e. $z = -\Delta z$) is taken to be unity. It can be shown that the scattering amplitude, to first order, is given by [24]

$$a(\mathbf{s}, \mathbf{s}_0) = k^2 (2\pi)^2 \exp(-k|s_{0z}|\Delta z) \int d\mathbf{K}_z \tilde{\eta}[\mathbf{k}(\mathbf{s}_\perp - \mathbf{s}_{\perp 0}), \mathbf{K}_z] I_h(\mathbf{K}_z; s_z, |s_{0z}|), \quad (14)$$

where

$$I_h(\mathbf{K}_z; s_z, |s_{0z}|) = \frac{\exp\{[k|s_{0z}| + i(ks_z - K_z)]d\}}{k|s_{0z}| + i(ks_z - K_z)}. \quad (15)$$

There is no longer a one-to-one correspondence between the scattering amplitude and the three-dimensional Fourier transform of the dielectric susceptibility, as there was in the case of conventional imaging [cf. Eq. (7)]. Specifically, the scattering amplitude $a(\mathbf{s}, \mathbf{s}_0)$ is now proportional to the weighted projection (i.e. the generalized Radon transform) along the K_z axis of all of those Fourier components of the susceptibility that have $\mathbf{K}_\perp = \mathbf{k}(\mathbf{s}_\perp - \mathbf{s}_{0\perp})$. The weighting function I_h , which is independent of object thickness, determines the effective number of longitudinal Fourier components that contribute to the scattering amplitude. Its normalized modulus, shown in Fig. 8, is peaked at $K_z = ks_z$ and has a nominal width of $2\sqrt{3}k|s_{0z}|$.

Figure 8. The modulus of the normalized weighting function, plotted for $|\mathbf{s}_{0z}| = 1.5$ and $s_z = 1$.

As a result, the scattered field in the direction \mathbf{s} carries information about a single high-frequency transverse Fourier component and many (both low and high frequency) longitudinal Fourier components, with the number of longitudinal Fourier components that effectively contribute increasing with the degree of evanescence $|\mathbf{s}_{0\perp}| = \sqrt{1 + |\mathbf{s}_{0z}|^2}$ of the incident field.

By contrast, for a two-dimensional object with susceptibility $\eta(\mathbf{r}) = \beta(\boldsymbol{\rho})\delta(z)$, the scattering amplitude takes the form

$$a(\mathbf{s}, \mathbf{s}_0) = k^2 (2\pi)^3 \exp(-k|\mathbf{s}_{0z}|\Delta z) \tilde{\beta}[\mathbf{k}(\mathbf{s}_{\perp} - \mathbf{s}_{0\perp})], \quad (16)$$

where

$$\tilde{\beta}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^2} \int_{z=0} d^2\rho \beta(\boldsymbol{\rho}) \exp(-i\boldsymbol{\xi} \cdot \boldsymbol{\rho}). \quad (17)$$

is the two-dimensional Fourier transform of $\beta(\boldsymbol{\rho})$. In this case, there is a one-to-one mapping between the high-frequency two-dimensional Fourier components of the object and the scattering amplitude. This one-to-one mapping explains the success of NSOM in resolving sub-wavelength detail in two-dimensional samples.

5. Image Reconstruction

As we have seen, for evanescent wave illumination and far-field detection, there is a *many-to-one* mapping between the three-dimensional spatial Fourier components of

the dielectric susceptibility and the scattered field. This leads to an inverse problem that is inherently ill-posed (or ill-conditioned in the discrete case) and generally under-determined [21]. In TIRT, there are two general reconstruction strategies that can be employed for three-dimensional structure determination: Fourier domain sampling and singular value decomposition. In either case, the inversion procedure must involve some type of regularization to deal with the problem of limited and noisy scattered field data. We will now discuss these two reconstruction strategies.

5.1 Fourier Domain Sampling

Fourier domain sampling is a procedure by which the Fourier components of the susceptibility are determined by discrete inversion of the generalized Radon transform represented by equation (14) [24, 25]. This sampling is typically achieved by one of two measurement schemes. The first scheme involves sampling the Fourier transform of the susceptibility $\tilde{\eta}(\mathbf{K})$ (for a given object orientation) by independent variation of \mathbf{s}_\perp and $\mathbf{s}_{0\perp}$, such that $|\mathbf{s}_\perp|$ remains fixed and $|\mathbf{s}_{0\perp}|$ assumes the discrete values $|\mathbf{s}_{0\perp}|^{(i)} = |\mathbf{s}_{0\perp}|^{(i-1)} + 2|\mathbf{s}_\perp|$, ($i = 1, 2, \dots, N$). A contiguous set of annular projection data is obtained, with each annular projection having the same width $2k|\mathbf{s}_\perp|$ but a different weighting $I_h^{(i)}$. This is illustrated in Fig. 9.

Figure 9. Illustrating two contiguous annular projection regions in the Fourier space of the dielectric susceptibility. The weighting of the projection data in each of the regions is clearly indicated. In addition, several typical projections in the

first region are shown, which correspond to the color-coded measurement directions illustrated in Fig. 7.

It follows that the inner and outer radii of the composite annulus are $K_{\min} = k \left[|\mathbf{s}_{0\perp}|^{(1)} - |\mathbf{s}_{\perp}| \right]$ and $K_{\max} = k \left[|\mathbf{s}_{0\perp}|^{(N)} + |\mathbf{s}_{\perp}| \right]$, respectively. If the measurement procedure is then repeated for all possible object orientations, a complete multi-view set of annular projection data is obtained that can be inverted for the three-dimensional Fourier transform of the susceptibility. As an example, for $|\mathbf{s}_{0\perp}|^{(1)} = 1.2$, $|\mathbf{s}_{0\perp}|^{(2)} = 2$, and $|\mathbf{s}_{\perp}| = 0.4$, we obtain two annular projections that cover the spatial frequency range $0.8 \leq K \leq 2.4$. Alternatively, we could achieve the same Fourier coverage by choosing $|\mathbf{s}_{0\perp}| = 1.6$ and $|\mathbf{s}_{\perp}| = 0.8$. Ultimately, the choice depends upon measurement noise and the availability of a given measurement.

Since $|\mathbf{s}_{\perp}| \leq 1$ and $|\mathbf{s}_{0\perp}|^{(i)} > 1$, there will be a spherical region of radius K_{\min} , centered about the origin, about which no Fourier information is available. This is not a problem in TIRT, since $K_{\min} \leq 2k$, and the low-frequency Fourier components can be determined from conventional scattering data [14]. The upper limit K_{\max} , which ultimately determines the spatial resolution, is a function of the distance between the exit face of the prism and the object, the thickness of the object, and the measurement noise. For TIRT, the theoretical maximum is $K_{\max} = k(n+1)$, so that the achievable spatial

resolution is limited to $\lambda/(n+1)$. For a prism with index of refraction $n = 2.4$, this yields a resolution limit of $\lambda/3.4$.

Alternatively, one can attempt a reconstruction for a fixed object orientation (i.e. single view). By this scheme, we fix $\xi = \mathbf{s}_{0\perp} - \mathbf{s}_{\perp}$ and vary $\mathbf{s}_{0\perp}$ and \mathbf{s}_{\perp} independently. We obtain a linear system of equations with the same three-dimensional Fourier components sampled by weighting functions of varying widths, which can be inverted. We then choose another ξ and repeat the process. The stability of the inversion depends upon the diversity one can achieve in the variation of $\mathbf{s}_{0\perp}$ and \mathbf{s}_{\perp} .

5.2 Singular-Value Decomposition

The second reconstruction strategy is based upon the singular value decomposition of the linearized scattering kernel [26-28]. It follows from Sec. 2 that, for a weakly scattering medium, the scattering amplitude (i.e. the data function) has the general linear form

$$a(\mathbf{s}_{\perp}, \mathbf{s}_{0\perp}) = \int \mathbf{K}(\mathbf{s}_{\perp}, \mathbf{s}_{0\perp}; \mathbf{r}) \eta(\mathbf{r}) d^3\mathbf{r} \quad (18)$$

or, in operator notation, $a = \mathbf{K}\eta$. The Fourier space analysis leads us to conclude that there exists a great deal of redundancy in the accessible data space. It is clear that a low-pass version of the susceptibility (with sub-wavelength detail) may be constructed from any number of subsets of the data. To obtain the best possible reconstruction, it is therefore desirable to take some linear combination of the available data to find a best solution. It is sensible to take as the best solution that solution which minimizes the

squared discrepancy between the forward modal (i.e. operator) acting on the solution and the actual data. That is, a solution that makes

$$\left\| \mathbf{K} \eta^+ - \mathbf{a} \right\|^2 \quad (19)$$

a minimum. In the event that such a solution is not unique, the solution of minimum norm is chosen. The operator \mathbf{K}^+ that connects the data to the solution is known as the pseudo-inverse:

$$\eta^+ = \mathbf{K}^+ \mathbf{a}. \quad (20)$$

To construct the operator \mathbf{K}^+ explicitly, the singular value decomposition (SVD) may be employed [21]. This approach offers several advantages, namely that the structure of the linear transformation is readily apparent, the effective degrees of freedom may be observed, and many regularization methods may be implemented by modification of the spectrum that is obtained. The SVD of the kernel \mathbf{K} is given by [28]

$$\mathbf{K}(\mathbf{s}_\perp, \mathbf{s}_{0\perp}; \mathbf{r}) = \sum_n \sigma_n f_n^*(\mathbf{r}) g_n(\mathbf{s}_\perp, \mathbf{s}_{0\perp}), \quad (21)$$

where σ_n is the singular value associated with the singular functions f_n and g_n . The $\{f_n\}$ and $\{g_n\}$ are orthonormal bases of the object and image Hilbert spaces, respectively, and are eigenfunctions with eigenvalues σ_n^2 of the positive self-adjoint operators $\mathbf{K}^+ \mathbf{K}$ and $\mathbf{K} \mathbf{K}^+$:

$$\mathbf{K}^+ \mathbf{K} f_n = \sigma_n^2 f_n \quad (22a)$$

$$\mathbf{K} \mathbf{K}^+ g_n = \sigma_n^2 g_n. \quad (22b)$$

In addition, the f_n and g_n are related by

$$\mathbf{K} f_n = \sigma_n g_n \quad (23a)$$

$$\mathbf{K}^+ g_n = \sigma_n f_n. \quad (23b)$$

For a band-limited (i.e. physical) operator, the singular values are a monotonically decreasing function of the index n , as shown in Fig. 10 [21].

Figure 10. Illustrating the behavior of the singular values for a band-limited operator.

It is interesting to note that the singular value decomposition is the functional equivalent of eigenfunction decomposition, except that the orthonormal basis functions f_n and g_n for the object space and image space, respectively, are different.

The kernel of the pseudo-inverse is readily obtained by the expression

$$\mathbf{K}^+(\mathbf{r}; \mathbf{s}_\perp, \mathbf{s}_{0\perp}) = \sum_n \frac{1}{\sigma_n} f_n(\mathbf{r}) g_n^*(\mathbf{s}_\perp, \mathbf{s}_{0\perp}), \quad (24)$$

where the sum is carried out over the n such that $\sigma_n \neq 0$. In the event that a true inverse exists, this pseudo-inverse reduces identically to it. When a true inverse does not exist, the pseudo-inverse yields the solution that minimizes the error on the orthogonal complement of the null space of \mathbf{K} .

In practice, small eigenvalues σ_n in Eq. (24) lead to unstable reconstructions [21]. In order to regularize the results and effectively deal with noisy data, the pseudo-inverse may be modified in a number of ways. By the Tikhonov method, the inverse

singular values $1/\sigma_n$ are replaced with $\sigma_n/(\sigma_n^2 + \beta^2)$, β being a tunable parameter. The modification of the spectrum is equivalent to solving a modified least squared error problem where in addition to minimizing the error, the quantity $\|\beta\eta^+\|^2$ is minimized simultaneously. When the spacing of the singular values is large compared to β , this method produces results very similar to a simple cut-off in the spectrum, i.e. a truncation in the sum over singular values to eliminate terms where $\sigma_n \leq \sigma_c$, where σ_c represents the cut-off singular value. Regularization effectively imposes a band limit on the reconstructions and so connects the noise level to the resolution.

6. Numerical Simulation

To illustrate the utility and power of TIRT, let us consider the following two-dimensional computer simulation. We take two point scatterers of diameter $d = \lambda/50$ to be separated along the x-axis by $\lambda/4$ and along the z-axis by $\lambda/10$ (see Fig. 11).

Figure 11. Illustrating the object used for the numerical simulation.

The distance along the z-axis between the prism face and the first scatterer is taken to be $\lambda/4$. The index of refraction of the prism is $n = 2.4$. We take 21 equally spaced angles of illumination in the backward half-space and 21 equally spaced angles of detection in the forward half-space. This corresponds to the following ranges for the incident and scattered transverse wave-vectors ($\mathbf{k}_{0\perp} = k\mathbf{s}_{0\perp}$ and $\mathbf{k}_{\perp} = k\mathbf{s}_{\perp}$):

$$\begin{aligned}
0 \leq |\mathbf{k}_{0\perp}| &\leq 2.4k \\
0 \leq |\mathbf{k}_{\perp}| &\leq k.
\end{aligned}
\tag{25}$$

In Figs. 12a and 12b, the transverse information content of the scattered field is illustrated for this case² and the case of homogeneous illumination (i.e. $n = 1$). Figure 12c represents the region of overlap between the two cases. We see that the use of evanescent incident waves increases the information content of the scattered field by roughly 70% with a corresponding increase in transverse resolution. But the real utility of the TIRT modality is the ability to resolve sub-wavelength features as a function of depth. This is clearly demonstrated in the following figures. Figure 13a illustrates the object reconstruction with no added noise (only machine error noise) and no regularization. The field of view of the figure is $\lambda/2$ by $\lambda/2$. Figure 13b illustrates the object reconstruction with 40 dB of additive noise that has been regularized. One can clearly resolve the two spheres. For purposes of comparison, Figure 13c illustrates the object reconstruction for a prism index of refraction of $n = 1$. This corresponds to the case that only homogeneous waves are used for illumination, and we see that it is difficult, even in this noise-free case to resolve the spheres.

Figure 13. Numerical simulation.

² In Fig. 11, we actually show the 2.5 – dimensional case (i.e. the case where the scattered field is measured in the entire forward half-space) for clarity. For the two-dimensional case, the bands would be lines along the x-axis.

7. Experimental Configuration

We are currently constructing a TIRT microscope at the NASA Glenn Research Center for use in biofluids research. Initially, the microscope will be 2.5 – dimensional, with a three-dimensional version to follow. The configuration of the microscope is shown in Fig. 14.

Figure 14. Illustrating the configuration of the prototype TIRT microscope.

The output from a frequency-doubled Nd: YAG laser is split into a sample beam (1%) and a reference beam (99%). The sample beam is scanned by a rotating mirror through an angular range $35^\circ \leq \theta_s \leq 65^\circ$ into a cylindrical glass prism of index of refraction $n = 1.9$. The critical angle of the prism is $\theta_c = 31.8^\circ$, so the scanned beam is totally internally reflected and evanescent waves are generated within the transverse wave vector range $1.09k \leq |\mathbf{k}_{0\perp}| \leq 1.72k$. The scattered light is collected by a high-quality objective with a numerical aperture (NA) of 0.9 (collection half-angle of 64.2°). This corresponds to a scattered field transverse wave-vector range of $0 \leq |\mathbf{k}_\perp| \leq 0.9k$. The incident and scattered wave-vector ranges are much more restricted in this case than in the numerical simulation due to practical limitations of the scanning and collection optics, respectively. Finally, the scattered light is heterodyned at a high-resolution CCD camera with the beam from the sample arm, which has been phase-modulated and expanded. Consequently, there is a direct mapping between the CCD output and the complex scattering amplitude over the aforementioned range. For each evanescent incident field, four CCD images

corresponding to four different reference beam phase shifts $(0, \pi/2, \pi, 3\pi/2)$ are taken to allow unambiguous determination of the phase of the complex scattering amplitude.

For comparison with the numerical simulation, Fig 15 illustrates the information content of the scattered field for the prototype TIRT microscope. We see the TIRT modality extends and complements the structural information that one would obtain with homogeneous incident fields alone.

Figure 15. Illustrating the transverse information content of the scattered field for the prototype TIRT microscope.

8. Conclusion

We have discussed the role of evanescent waves in achieving the super-resolution of near-field imaging modalities. The essential mechanism, the interaction of the super-oscillatory evanescent fields with the sub-wavelength structure of the sample, may be invoked by means other than the usual probe-induced field localization. In fact, a single evanescent plane wave may be generated at the surface of a prism and used as a source of illumination as is routinely done in TIRM. A careful analysis of the scattered field reveals that the Ewald sphere of reflection is shifted by the incident wave vector as is well understood in diffraction tomography. However, because the magnitude of the transverse part of the evanescent wave vector is larger than the free-space wave number, the region of the Fourier space of the object that is accessible now includes points outside the usual Ewald limiting sphere. For this reason, when data are collected for a range of incident

fields, the classical resolution limits may be overcome in a computed reconstruction of the sample susceptibility. This approach yields the additional benefit that the reconstruction is inherently tomographic and so a view of the sample structure as a function of depth may be obtained.

Acknowledgements: DGF would like to acknowledge the support of the NASA Microgravity Fluid Physics Program, as well as the contributions of Marius Asipauskas in the construction of the prototype TIRT microscope. PSC would like to acknowledge support from the National Aeronautics and Space Administration under Grant NAG3-2764 and the National Science Foundation under CAREER grant 0239265.

References

1. Betzig, E. and J.K. Trautman, *Near-Field Optics - Microscopy, Spectroscopy, and Surface Modification Beyond the Diffraction Limit*. Science, 1992. **257**(5067): pp. 189-195.
2. Durig, U., D.W. Pohl, and F. Rohner, *Near-Field Optical-Scanning Microscopy*. Journal of Applied Physics, 1986. **59**(10): pp. 3318-3327.
3. Courjon, D. and C. Bainier, *Near-Field Microscopy and near-Field Optics*. Reports on Progress in Physics, 1994. **57**(10): pp. 989-1028.
4. Girard, C. and A. Dereux, *Near-field optics theories*. Reports on Progress in Physics, 1996. **59**(5): pp. 657-699.
5. Temple, P.A., *Total Internal-Reflection Microscopy - a Surface Inspection Technique*. Applied Optics, 1981. **20**(15): pp. 2656-2664.
6. Axelrod, D., *Total internal reflection fluorescence microscopy in cell biology*. Traffic, 2001. **2**(11): pp. 764-774.
7. Stock, K., et al., *Variable-angle total internal reflection fluorescence microscopy (VA-TIRFM): realization and application of a compact illumination device*. Journal of Microscopy-Oxford, 2003. **211**: pp. 19-29.
8. Sarayeddine, K., et al., *Scanning Tunneling Optical Microscopy - a New Tool for High-Resolution Nondestructive Testing*. Institute of Physics Conference Series, 1990(98): pp. 29-32.
9. Guerra, J.M., *Photon Tunneling Microscopy*. Applied Optics, 1990. **29**(26): pp. 3741-3752.
10. Courjon, D., K. Sarayeddine, and M. Spajer, *Scanning Tunneling Optical Microscopy*. Optics Communications, 1989. **71**(1-2): pp. 23-28.
11. Kowarz, M.W., *Homogeneous and Evanescent Contributions in Scalar near-Field Diffraction*. Applied Optics, 1995. **34**(17): pp. 3055-3063.
12. Blattner, P., H.P. Herzig, and R. Dandliker, *Scanning near-field optical microscopy: transfer function and resolution limit*. Optics Communications, 1998. **155**(4-6): pp. 245-250.
13. Born, M. and E. Wolf, *Principles of Optics*. 7 ed. 1999, Cambridge: Cambridge University Press.
14. Wolf, E., *Three-dimensional structure determination of semi-transparent objects from holographic data*. Optics Communications, 1969. **1**: pp. 153-156.

15. Abbe, E., *Archiv. f. Mikroskopische Anat.*, 1873. **9**: pp. 413.
16. Rayleigh, L., *Philosophical Magazine*, 1879. **8**: p. 261.
17. James, R.W., *The Optical Principles of the Diffraction of X-rays*. 1948, London: G. Bell & Sons.
18. Fuks, B.A., *Introduction to the Theory of Analytic Functions of Several Complex Variables*. 1963, Providence: American Mathematics Society.
19. Viano, G.A., *Journal of Mathematical Physics*, 1976. **17**: pp. 1160-1165.
20. Habashy, T. and E. Wolf, *Reconstruction of Scattering Potentials from Incomplete Data*. *Journal of Modern Optics*, 1994. **41**(9): pp. 1679-1685.
21. Bertero, M. and P. Boccacci, *An Introduction to Inverse Problems in Imaging*. 1998, Bristol: Institute of Physics.
22. Synge, E., *A suggested method for extending microscope resolution into the ultra-microscopic region*. *Philosophical Magazine*, 1928. **6**: pp. 356-362.
23. Ash, E. and G. Nicholls, *Super-resolution aperture scanning microscope*. *Nature*, 1972. **237**: pp. 510-512.
24. Fischer, D.G., *The information content of weakly scattered fields: implications for near-field imaging of three-dimensional structures*. *Journal of Modern Optics*, 2000. **47**(8): pp. 1359-1374.
25. Fischer, D.G., *Subwavelength depth resolution in near-field microscopy*. *Optics Letters*, 2000. **25**(20): pp. 1529-1531.
26. Carney, P.S. and J.C. Schotland, *Three-dimensional total internal reflection microscopy*. *Optics Letters*, 2001. **26**(14): pp. 1072-1074.
27. Carney, P.S. and J.C. Schotland, *Theory of total-internal-reflection tomography*. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 2003. **20**(3): pp. 542-547.
28. Carney, P.S. and J.C. Schotland, *Near-field tomography*, in *Inside Out: Inverse Problems and Applications*, G. Uhlmann and S. Levy, Editors. 2003, Cambridge University Press: Cambridge.

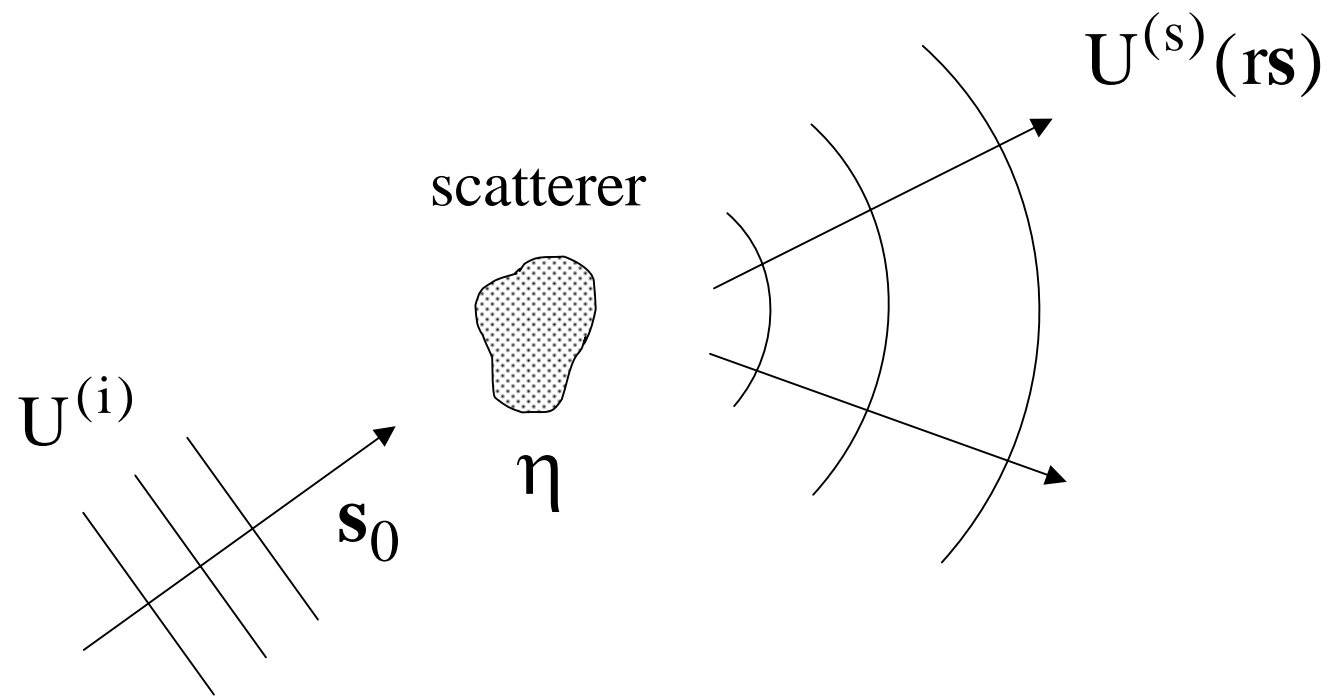


Figure 1

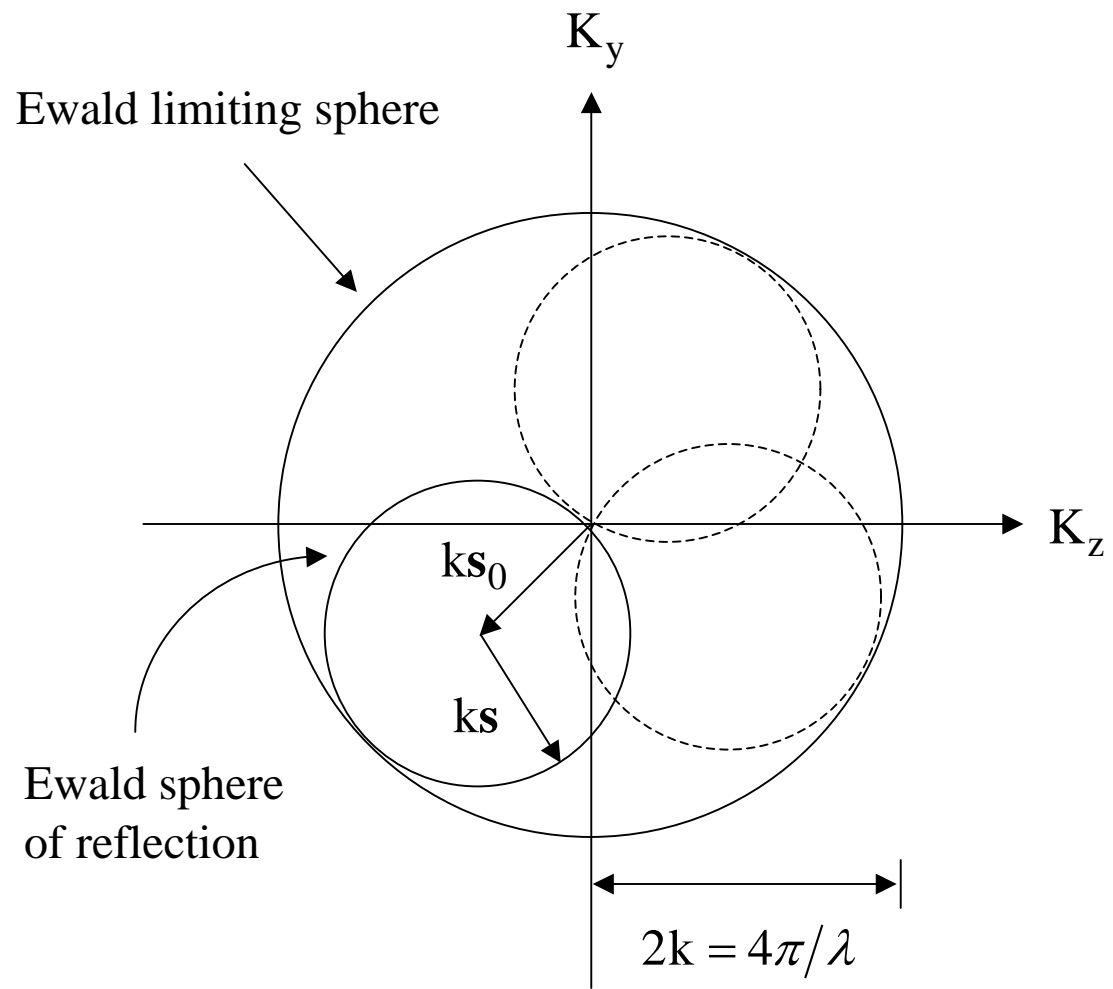
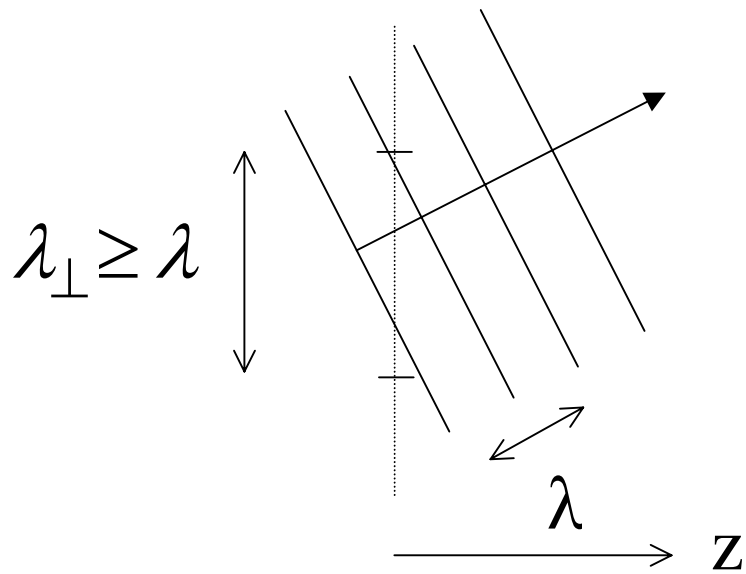


Figure 2

homogeneous wave



evanescent wave

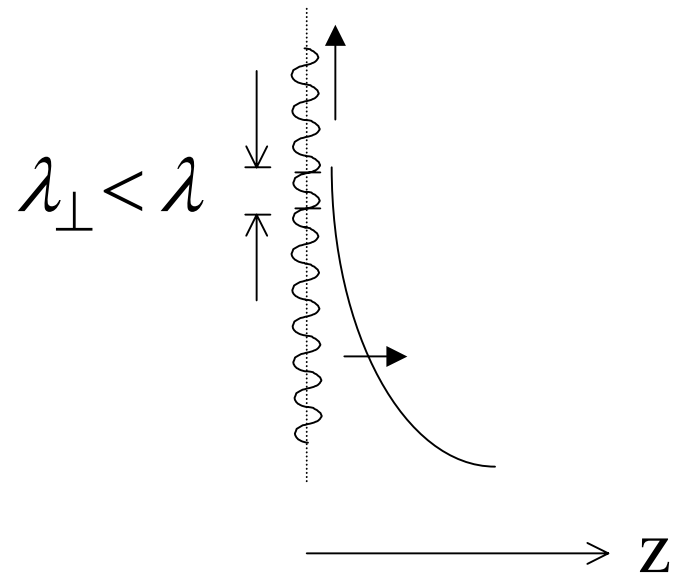


Figure 3

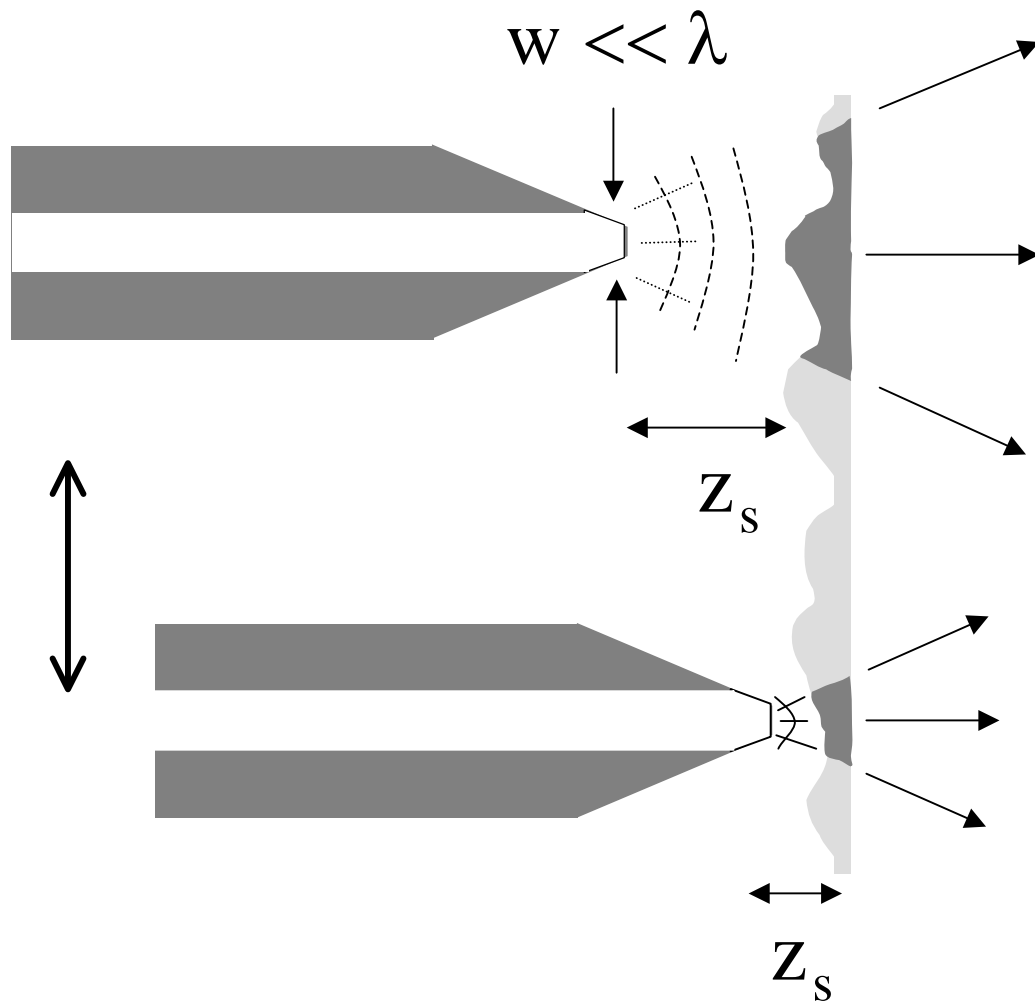


Figure 4

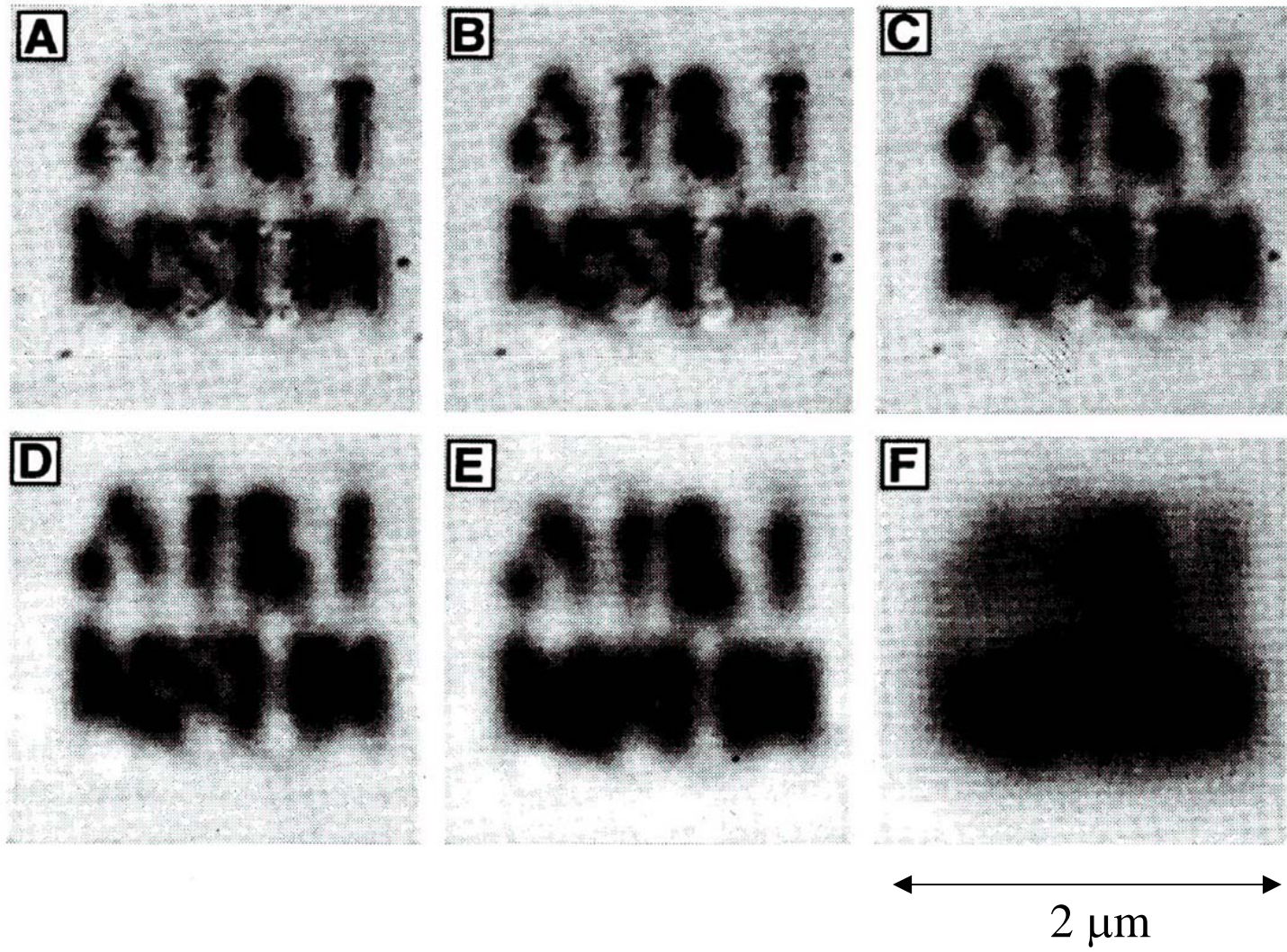
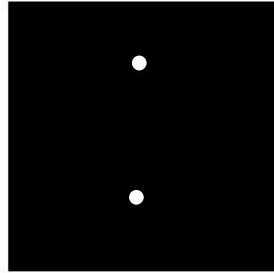
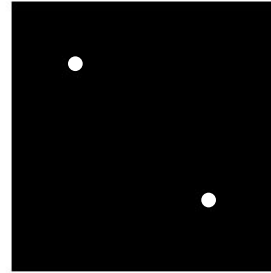


Figure 5

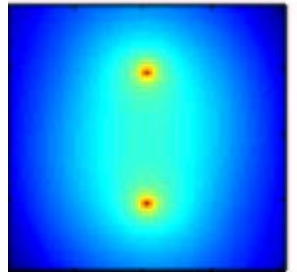
$z = 0.005\lambda$



$z = 0.405\lambda$

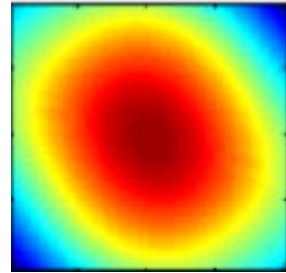


λ



λ

λ



λ

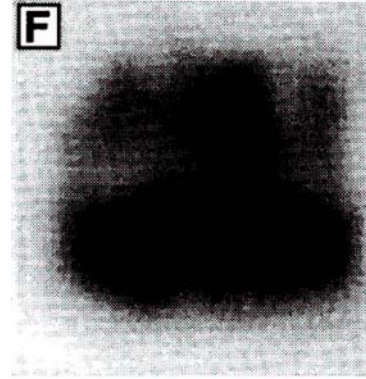
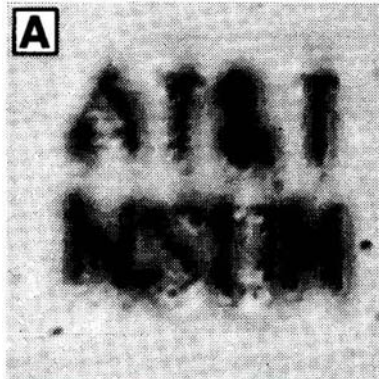


Figure 6

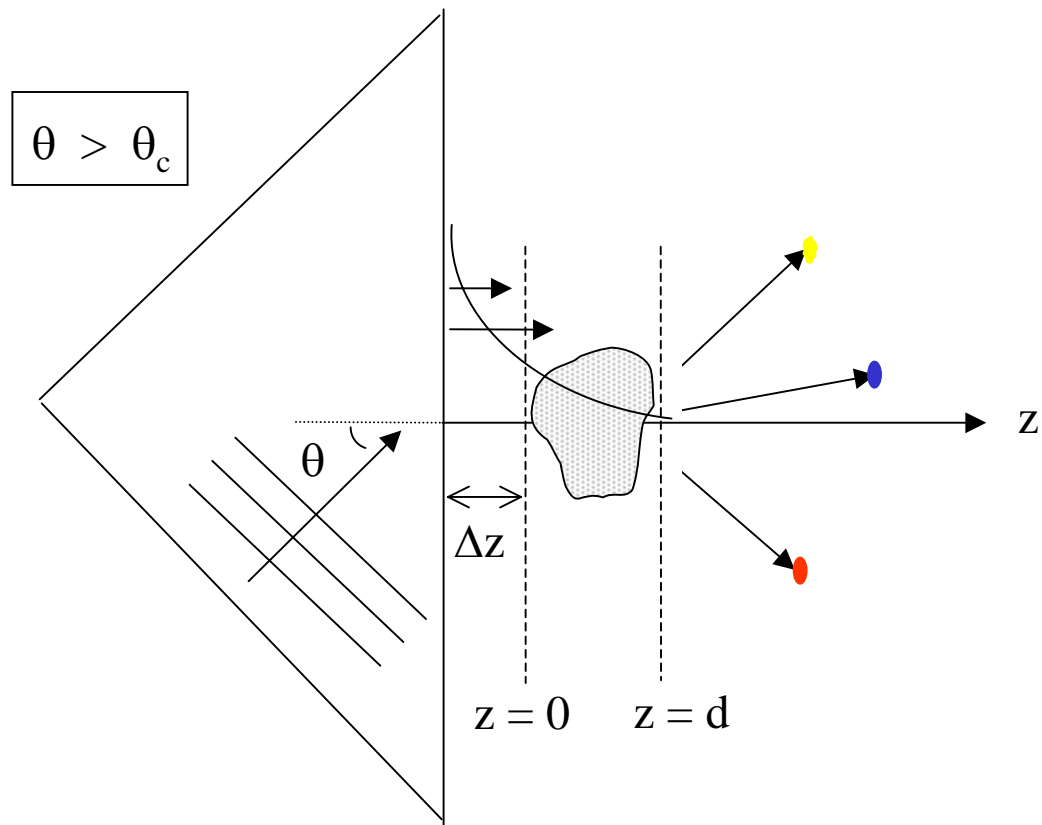


Figure 7

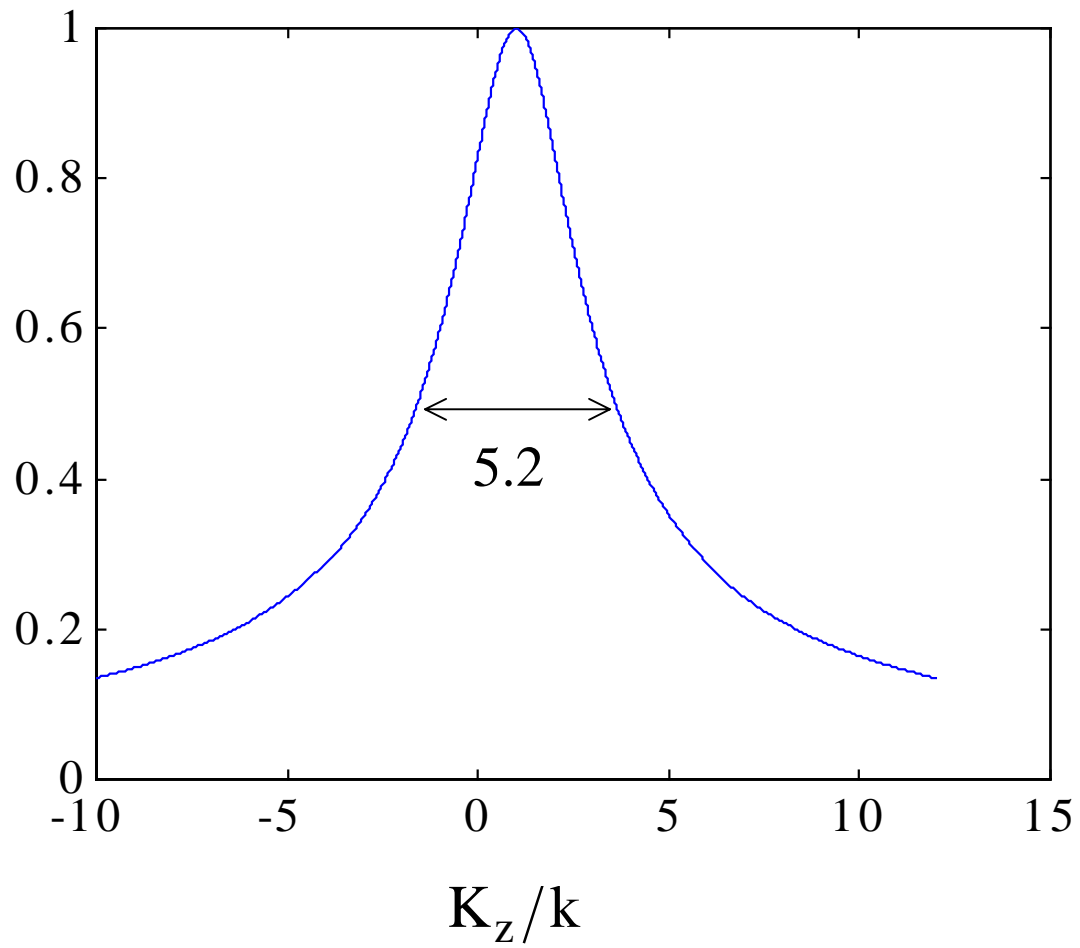


Figure 8

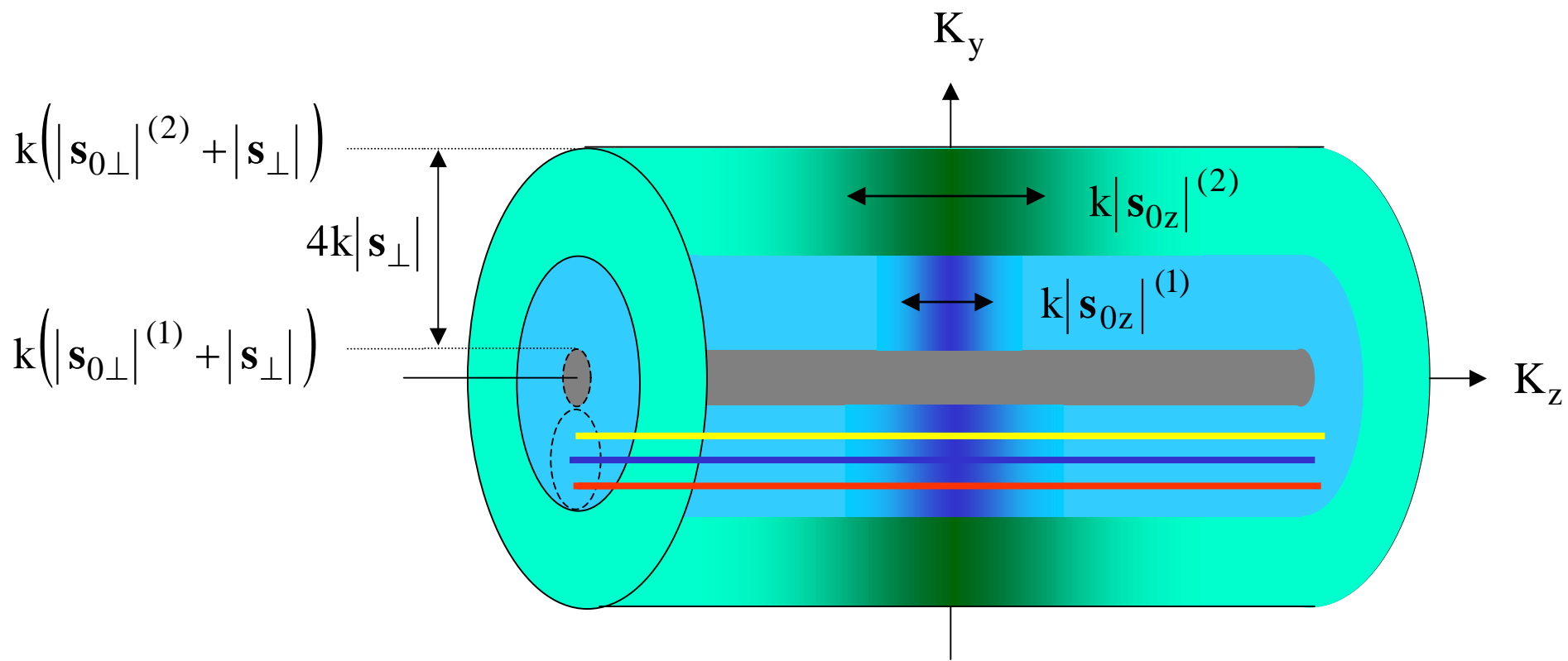


Figure 9

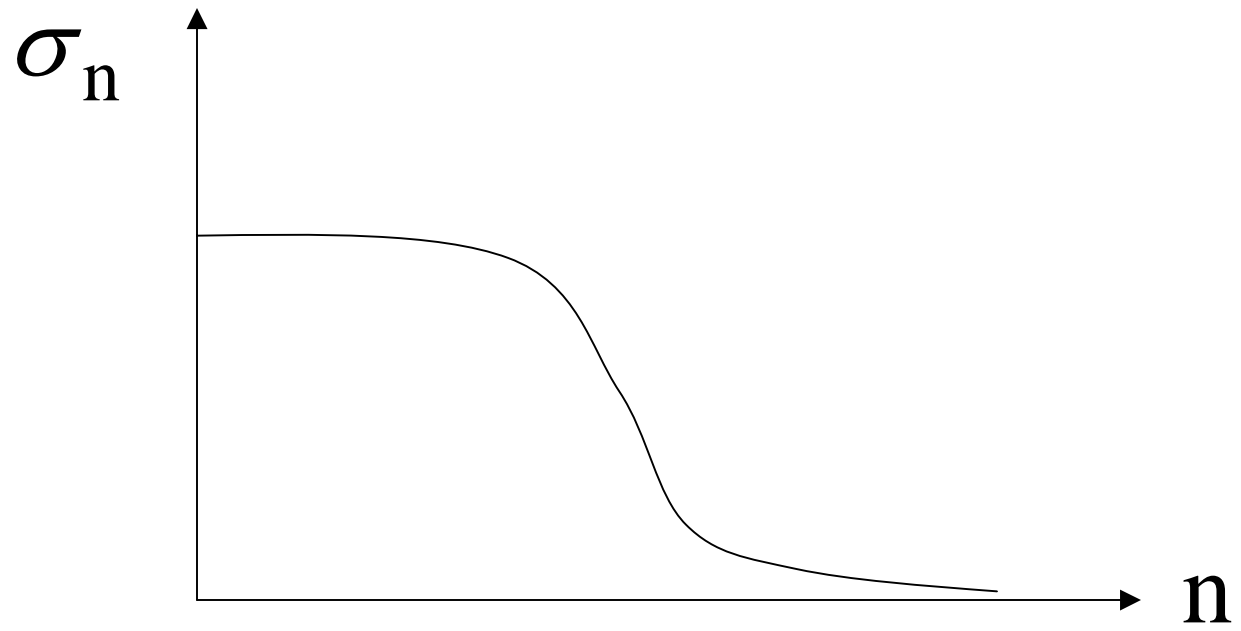


Figure 10

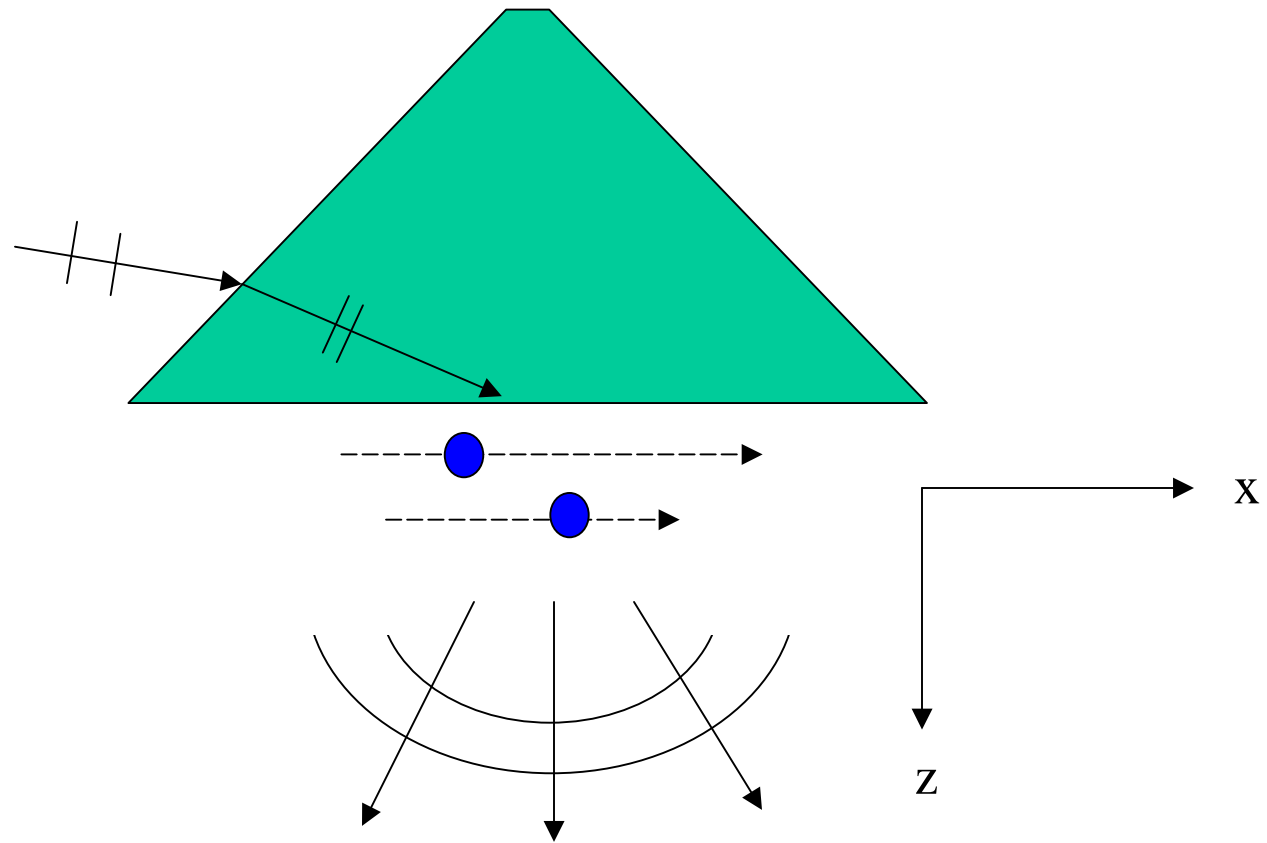


Figure 11

TIRT ($n = 2.4$)

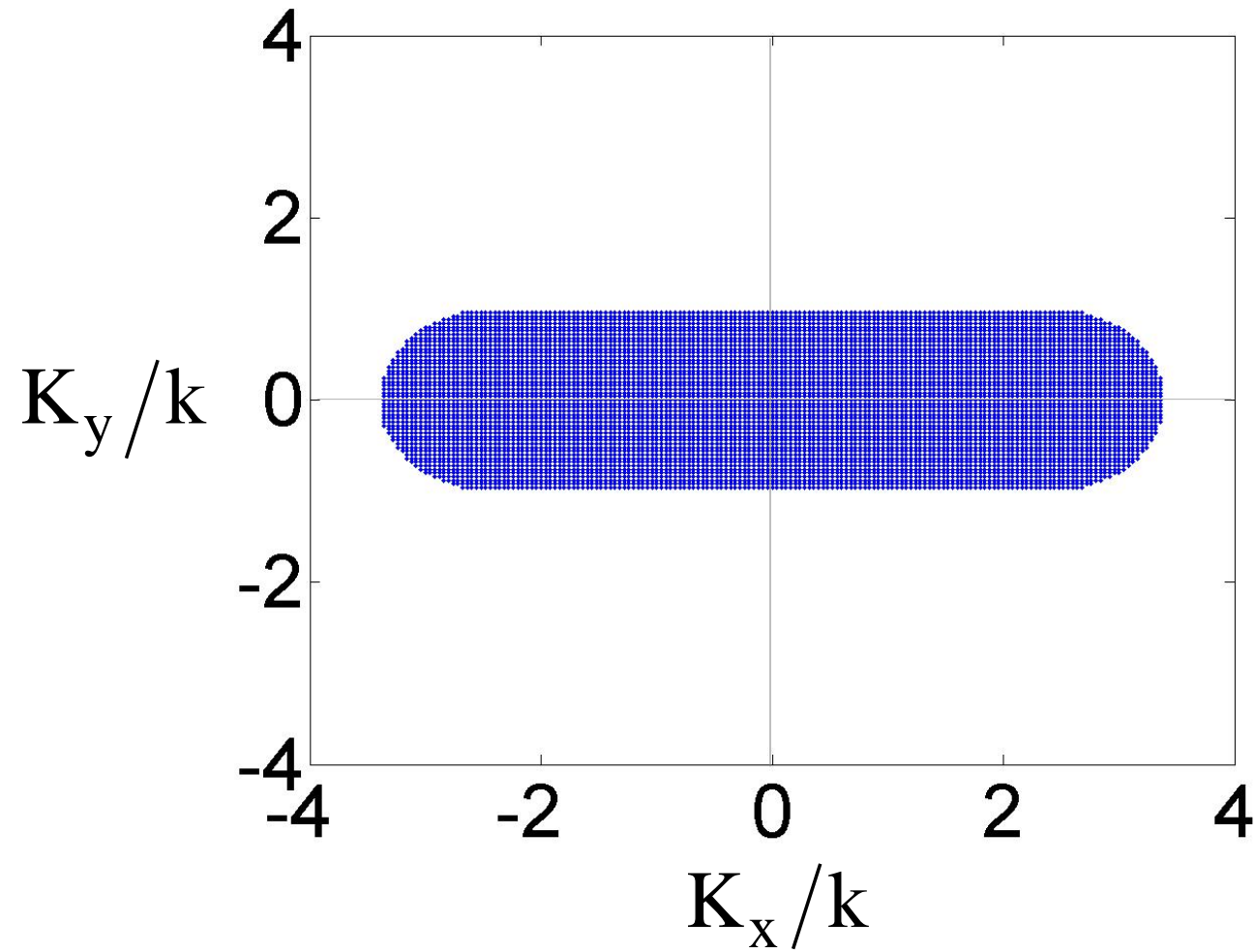


Figure 12a

Homogeneous Scattering ($n = 1$)

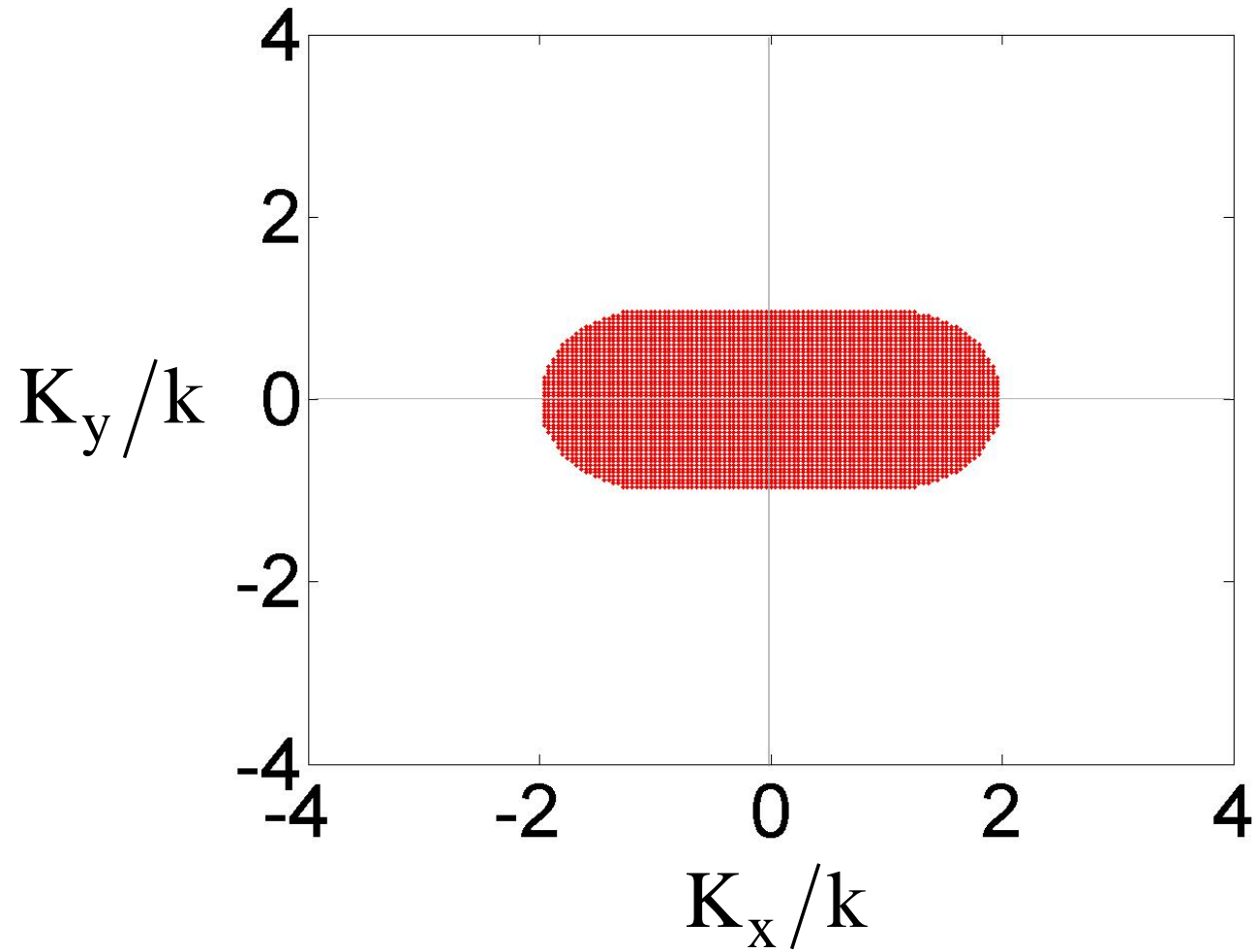


Figure 12b

Overlap Region

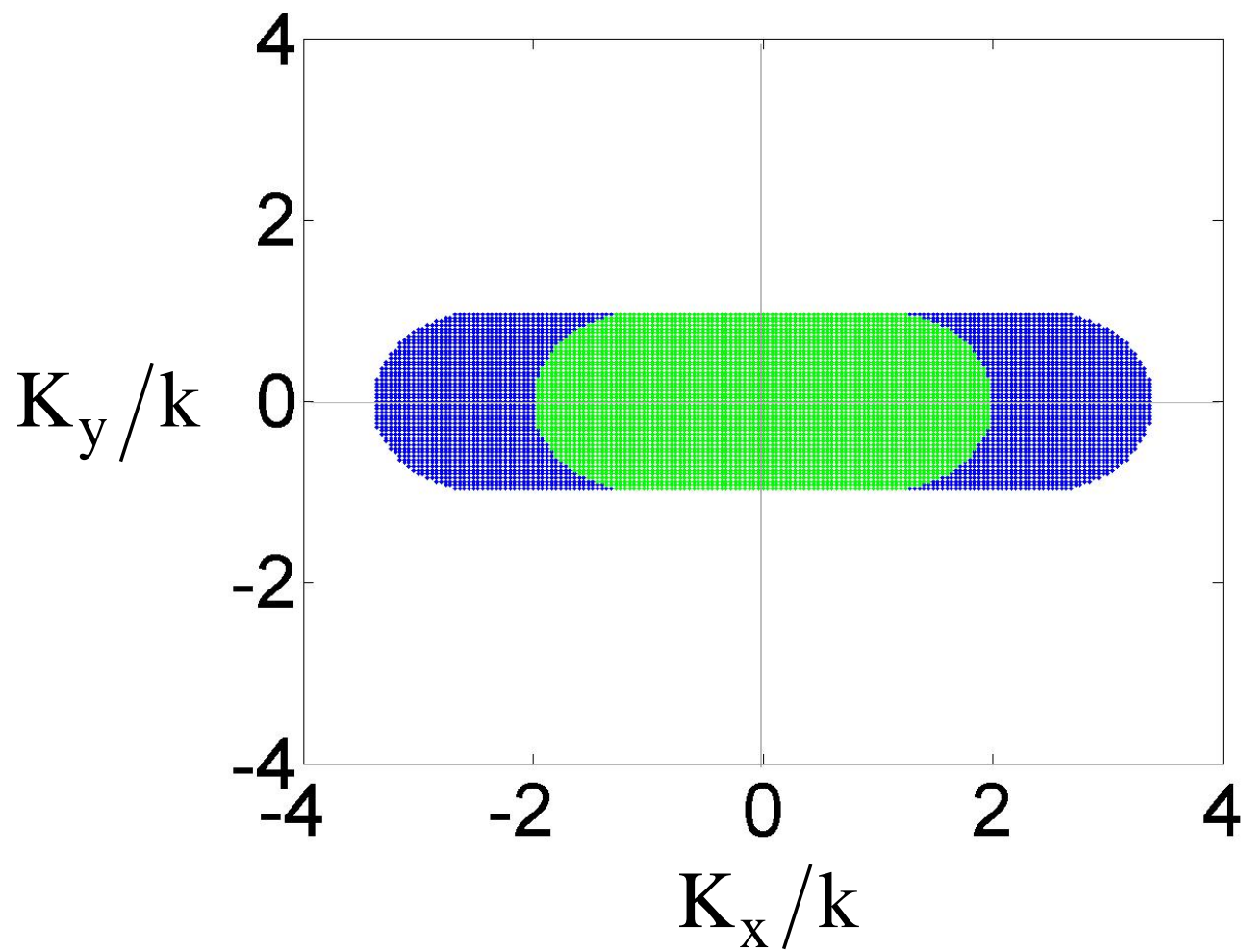


Figure 12c

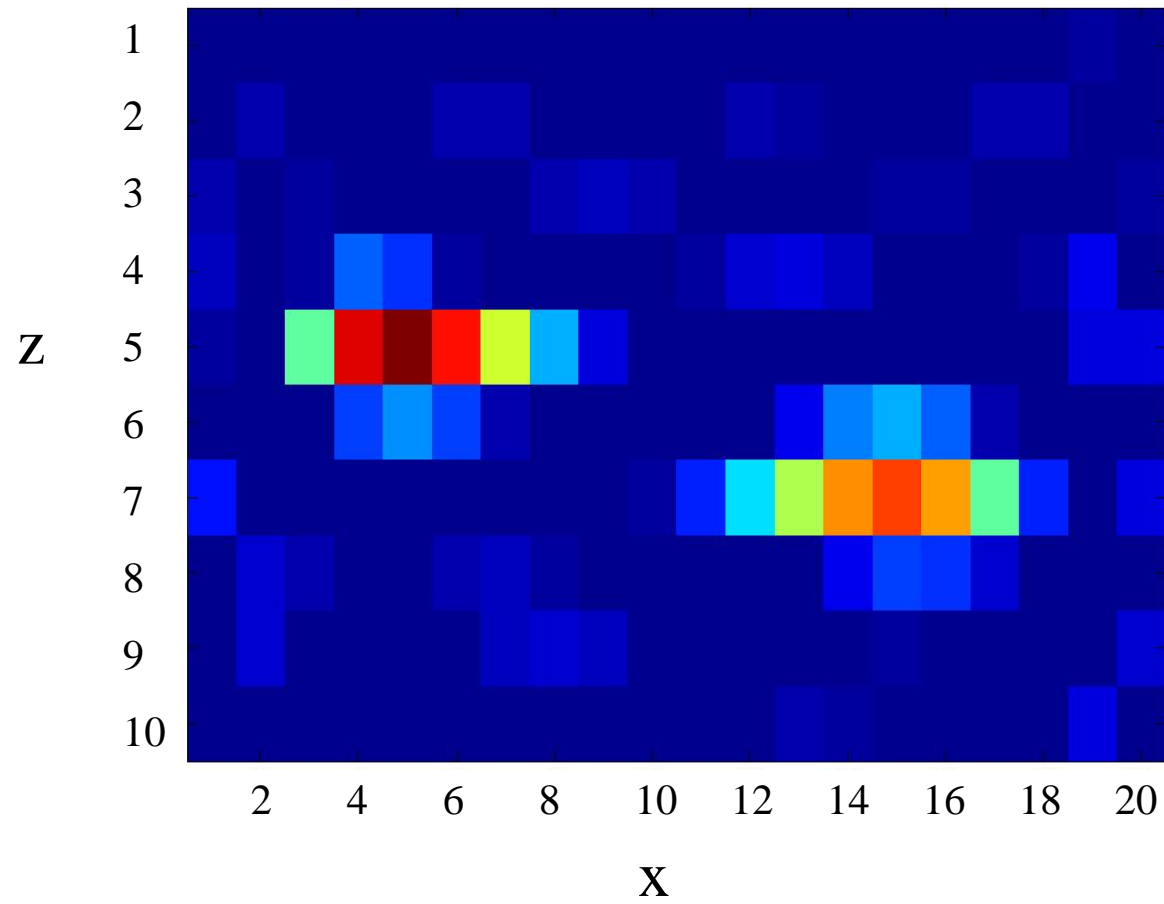


Figure 13a

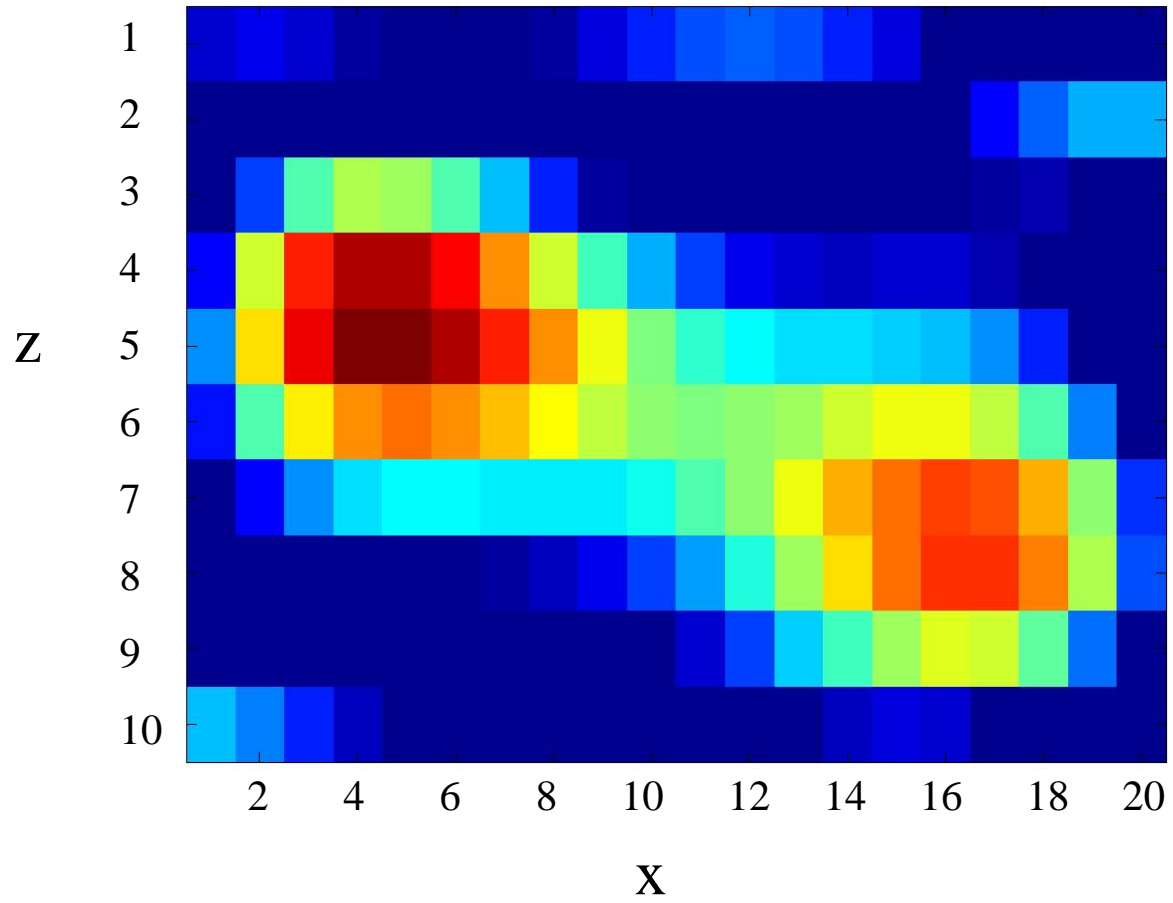


Figure 13b

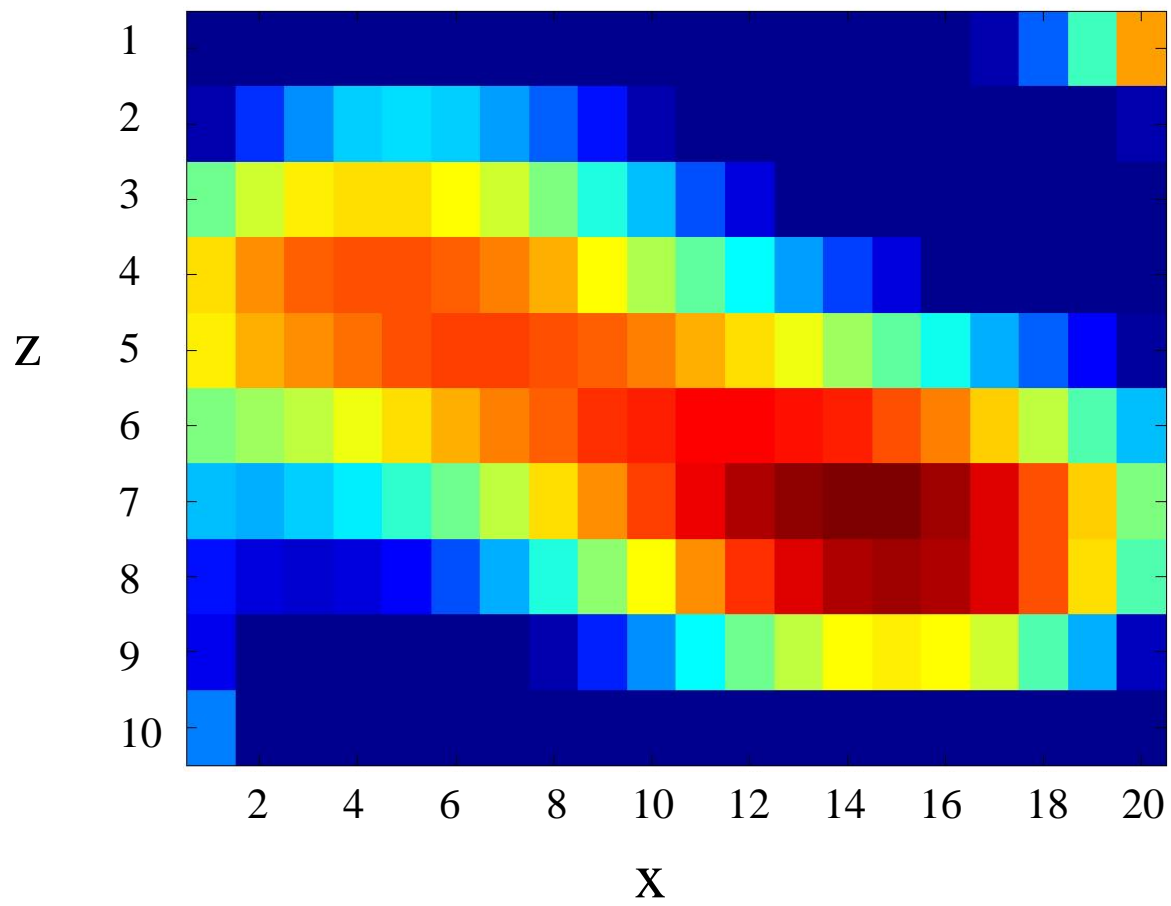


Figure 13c

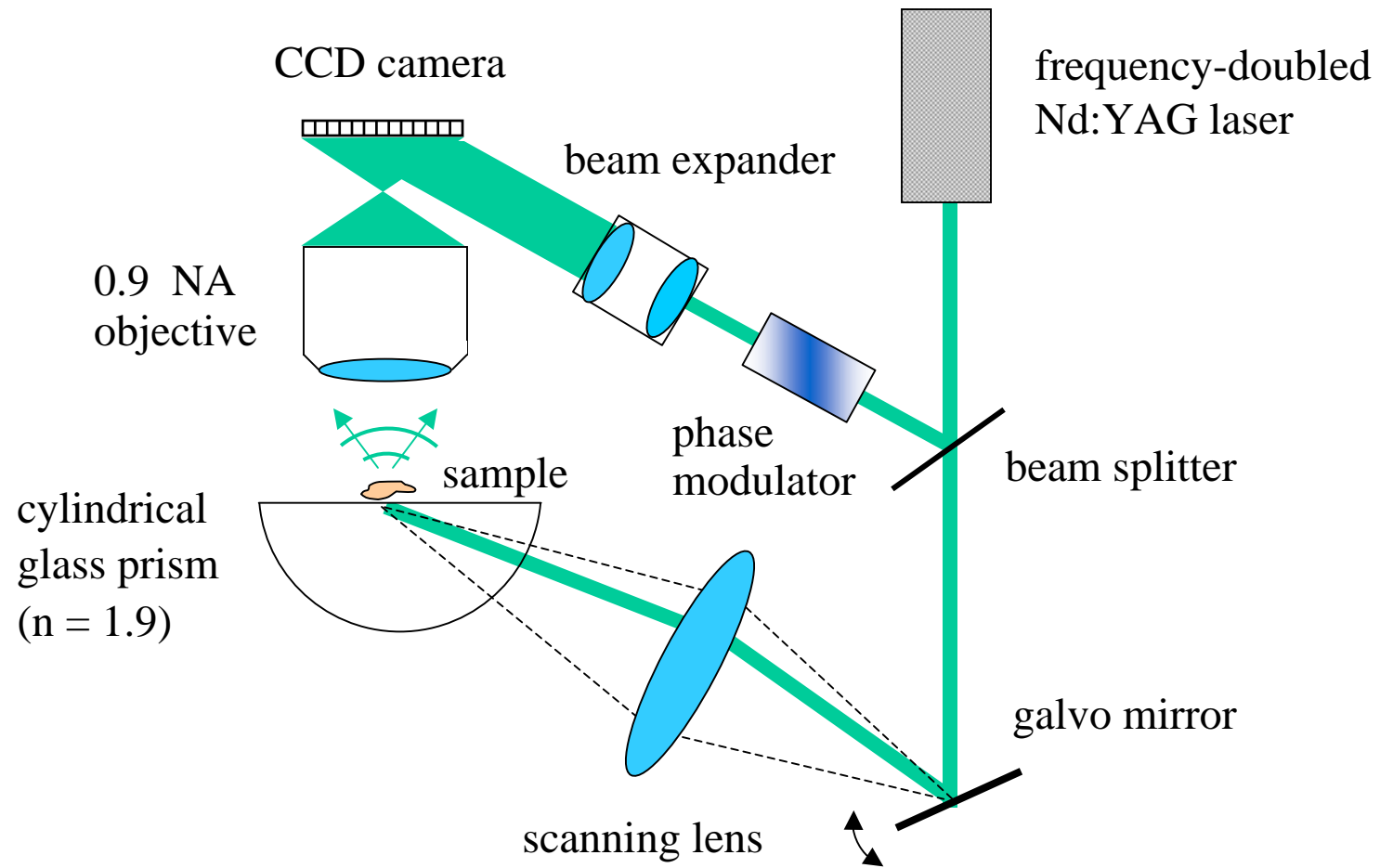


Figure 14

TIRT (n = 1.9)

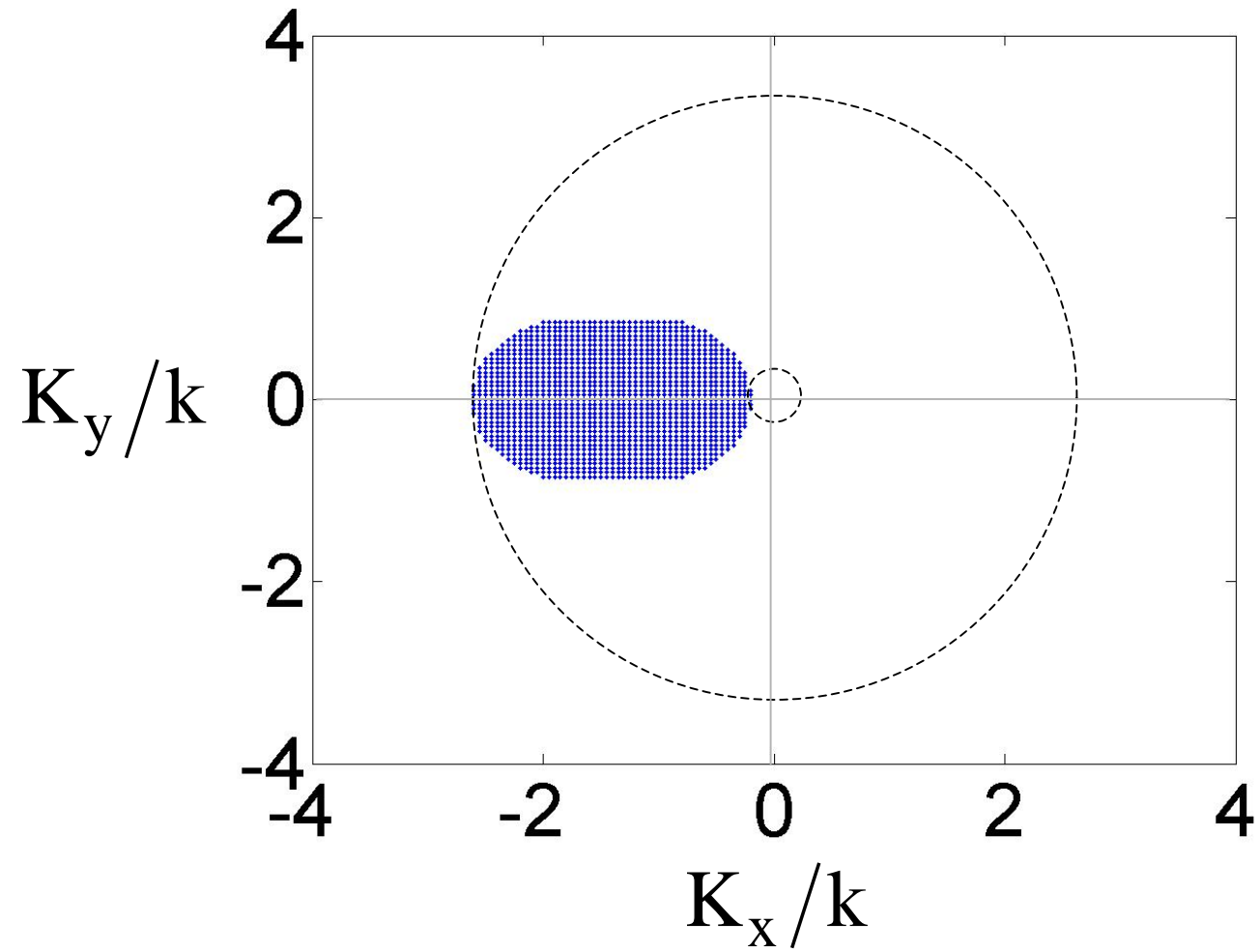


Figure 15a

Homogeneous Scattering ($n = 1$)

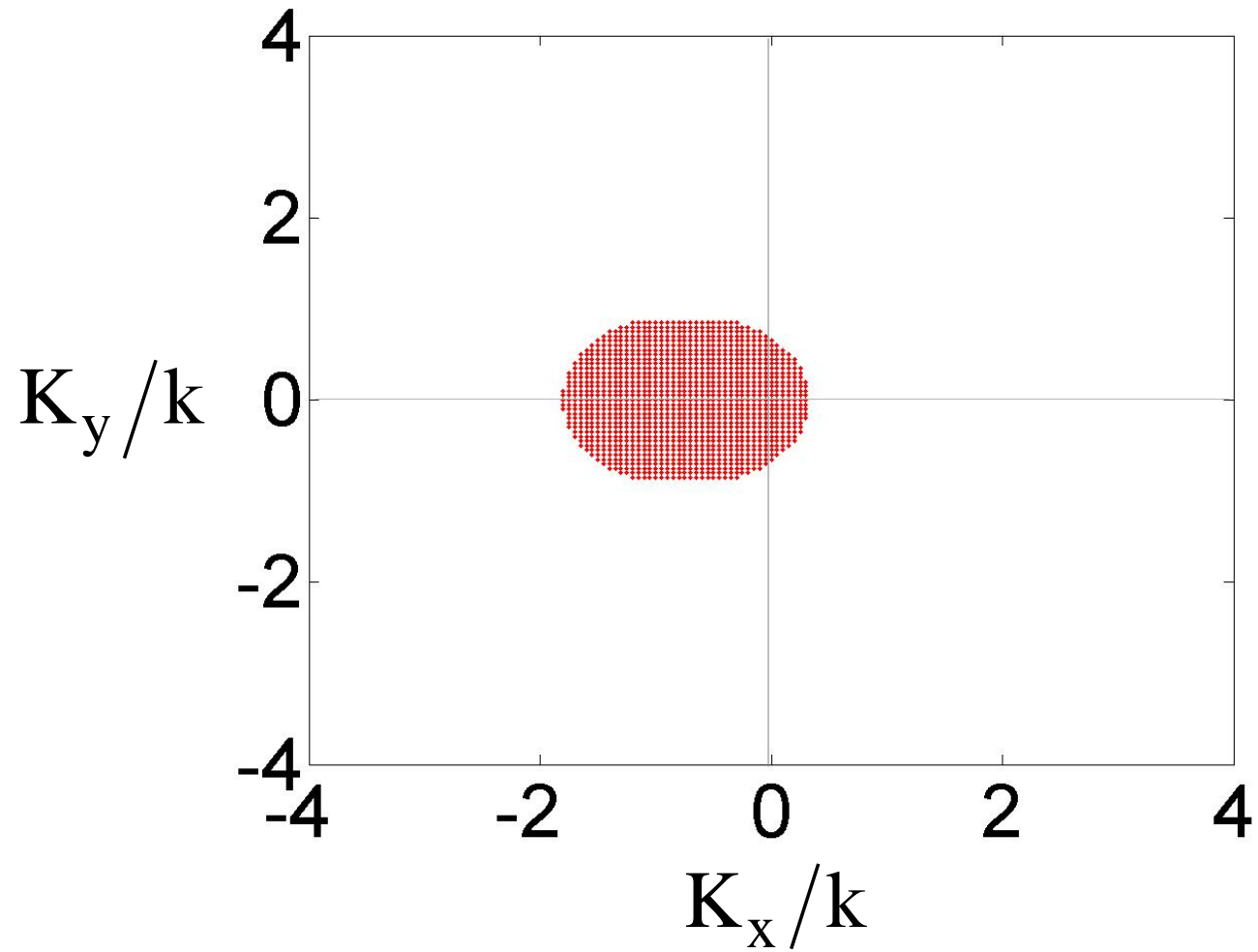


Figure 15b

Overlap Region

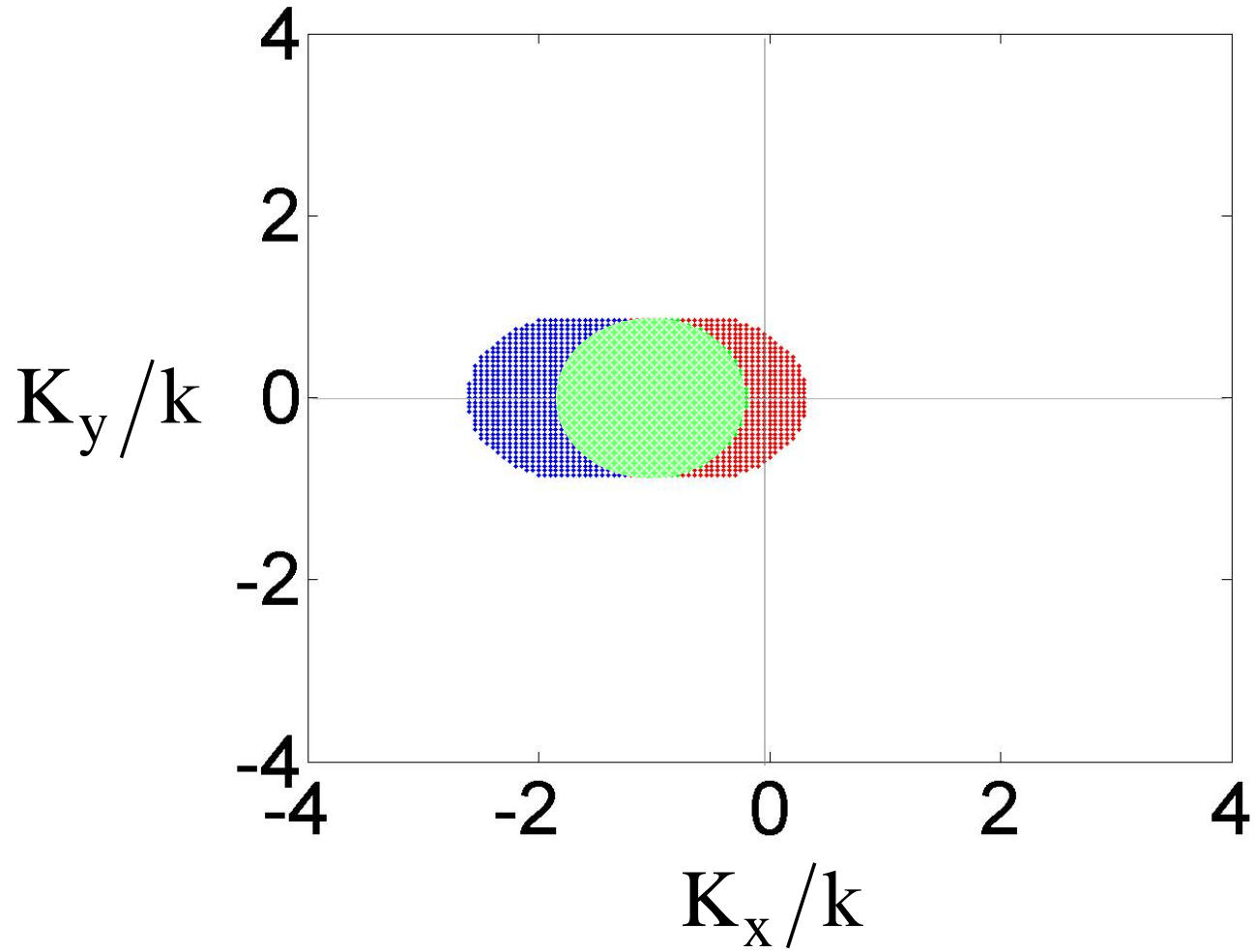


Figure 15c