Volume 1 of 2 (Text)

DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

BY

THOMAS JOSEPH CAVICCHI

B.S., Massachusetts Institute of Technology, 1982
M.S., University of Illinois, 1984

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1988

Urbana, Illinois

University of Illinois at Urbana-Champaign

DEPARTMENTAL FORMAT APPROVAL

THIS IS TO CERTIFY THAT THE FORMAT AND QUALITY OF PRESENTATION OF THE THESIS SUBMITTED BY THOMAS JOSEPH CAVICCHI AS ONE OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IS ACCEPTABLE TO THE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
(Full Name of Department, Division or Unit)

17 February 1988
Date of Approval

Richard Buhr
Departmental Representative

DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

BY

THOMAS JOSEPH CAVICCHI

B.S., Massachusetts Institute of Technology, 1982
M.S., University of Illinois, 1984

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1988

Urbana, Illinois

DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

BY

THOMAS JOSEPH CAVICCHI

B.S., Massachusetts Institute of Technology, 1982
M.S., University of Illinois, 1984

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1988

Urbana, Illinois

DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

BY

THOMAS JOSEPH CAVICCHI

B.S., Massachusetts Institute of Technology, 1982
M.S., University of Illinois, 1984

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1988

Urbana, Illinois

DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

BY

THOMAS JOSEPH CAVICCHI

B.S., Massachusetts Institute of Technology, 1982
M.S., University of Illinois, 1984

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1988

Urbana, Illinois

DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

BY

THOMAS JOSEPH CAVICCHI

B.S., Massachusetts Institute of Technology, 1982
M.S., University of Illinois, 1984

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1988

Urbana, Illinois

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

THE GRADUATE COLLEGE

FEBRUARY 1988

WE HEREBY RECOMMEND THAT THE THESIS BY

THOMAS JOSEPH CAVICCHI

ENTITLED DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

William D. O'Brien, Jr.
Director of Thesis Research

N. Narayana Rao
Head of Department

Committee on Final Examination†

William D. O'Brien, Jr.
Chairperson

W. Kenneth Jenkins

Richard G. Hay

† Required for doctor's degree but not for master's.

O-517

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

THE GRADUATE COLLEGE

FEBRUARY 1988

WE HEREBY RECOMMEND THAT THE THESIS BY

THOMAS JOSEPH CAVICCHI

ENTITLED DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

William D. O'Brien, Jr.
Director of Thesis Research

N. Narayana Rao
Head of Department

Committee on Final Examination†

William D. O'Brien, Jr.
Chairperson

W. Kenneth Jenkins

Richard G. Hay

† Required for doctor's degree but not for master's.

O-517

DIFFRACTION TOMOGRAPHY AND THE SINC BASIS MOMENT METHOD

Thomas Joseph Cavicchi, Ph.D.
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign, 1988

ABSTRACT

To obtain high-resolution quantitative images of acoustic parameters (such as sound speed and absorption coefficient) in tissue from measurements of the interaction of incident waves of ultrasound with the tissue, diffraction effects must be included. The vast majority of all diffraction tomography algorithms in existence depend on first-order scattering assumptions for their validity which do not hold in practical tissue inverse scattering problems. The present work investigates several aspects of a higher-order algorithm, the Sinc Basis Moment Method. First, conventional tomography algorithms, including those having straight-path and first-order scattering assumptions, are compared. Close attention is given to a comparison of the mathematical meaning of the first Born and Rytov approximations. Then, the equations of the higher-order sinc basis method are explained, and signal processing details for its implementation are given. The scatterer chosen for most of this work is the circular cylinder, for which exact scattered field data resulting from an incident cylindrical wave may be calculated independently of the reconstruction equations. Object parameters such as sound speed and absorption contrast and size, as well as algorithm parameters such as sampling density, grid size, and relaxation con-

stants, are varied to determine behavior and limitations of the algorithm. The algorithm itself was modified to use knowledge about the problem structure to maximize computational efficiency. In addition, an interesting use of the FFT which significantly reduces the order of computation is described. The first iteration of the Sinc Basis Method is shown to be equivalent to a typical first-order, Born-approximation-based solution. Finally, use of a minisupercomputer has helped make evident a fundamental limitation of the algorithm, the size of the phase shift of a wave passing through the object. An abrupt threshold of reconstruction quality exists near $\pm\pi$. The reason appears to be the existence of multiple solutions arising from the periodicity in phase representations. The iterative (perturbation) technique settles upon the closest solution to the starting point, resulting in erroneous reconstructions when the closest solution is no longer the desired solution (which is true for phase shift magnitudes greater than π). For the ambiguity to be resolved, either the initial starting points for both the object function and the field must be substantially improved, or somehow the information contained in unwrapped phase measurements must be preserved in the computations.

ACKNOWLEDGEMENTS

The topic of this thesis was introduced to me by my research advisor, Professor William D. O'Brien, Jr., and all along he has been very helpful in a variety of ways, most notably in various insights he has had about ideas and how best to arrange and prioritize them for presentation. He also was responsible for the funding, and was an especially vital source of patience and optimism.

Michael Haney has been another friend and advisor who had a bottomless pit of generosity of his time and wisdom that was for me a true mystery. In addition, his questions were of a nature that forced me to try to rise higher, both technically and philosophically.

Many others in the Bioacoustics Research Laboratory have helped out with technical matters in this thesis. Bob Cicone was tireless in offering his expertise on figures (he never made a mistake). Joe Cobb has been a friendly assistant with computer matters, and "brought things up" at crucial times. Several professors and graduate students have also helped me even when they did not know it--when they were willing to be sounding boards (with varying reflection and absorption coefficients) for my various verbal insonifications.

A special thanks to Nancy Dimond and Wanda Elliott for their endless energy in expertly executing my endless requests in word processing. Also thanks to Barbara Kaptain. The three people above were very friendly people to talk with about anything from

alpha to omega during spare moments (notice that I here avoided Greek symbols by spelling them out!).

Another round of applause for some of my very best friends in Champaign outside the laboratory who gave encouragement in my moments of despair and put up with my craziness during my better moods: Abraham Varghese, Tom Baughman, and Raphael Gely.

This thesis is humbly dedicated to my family: mother, father, brother and sister, who have shared in the best times of my life. How can anyone be adequately thanked who supports and gives so much in an infinity of ways over the span of a whole lifetime?

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.0 Introduction

A tomogram is an image of a cross section of a body. In the area of medical imaging it can be used for diagnosis if the value of the parameter displayed differs in pathological tissue from its value in normal tissue. B-scan images, which are images of the scattered field, only show structure such as interfaces; local information about tissue parameters is not directly available from B-scans. To improve correctness of diagnosis it is desirable to have such information to develop thresholds of pathology or signatures of pathological tissue. Computer tomography provides quantitative images of one or more relevant parameters. Currently, several source modalities are either in use or are being researched. Among these are microwave and ultrasonic tomography, which image electromagnetic and mechanical or attenuative parameters, respectively.

The focus of this study is on tomography for the case in which the sizes of the scatterers in the tissue are of the same order as the wavelength of the incident field. Under this condition, which holds for both ultrasonic and microwave tomography, diffraction effects are significant and the inhomogeneous wave equation must be inverted. If the scattering is not too strong, mathematical/numerical methods can be used to reconstruct the parameter(s) of interest.

This thesis begins with a short survey of some of the ultrasonic and microwave tomography literature relevant to the present work. Presented in Chapter 2 is a detailed exposition of several conventional tomography algorithms, including relationships between them, all presented in a notation unified as much as possible. Included in that chapter are discussions of the Born and Rytov approximations, which must be understood in order to assess the sinc basis moment method relative to other existing algorithms. In Chapter 3 the sinc basis method is derived and discussed, and the bulk of early computational results performed on the Bioacoustics Research Laboratory minicomputer, a VAX 11/730, is included. Also, a short discussion of the Algebraic Reconstruction Technique is given, and related to the use of this method in the present work. A derivation for the exact scattered field both inside and outside a circular cylindrical object insonified by an incident cylindrical wave is given, allowing for testing of the algorithm independent of the sinc basis reconstruction equations. A method devised to use the fast Fourier transform to reduce the order of computation of the slower of the two sets of matrix equations is given in Chapter 4, including both theoretical and graphical/computational arguments for its validity. Finally, in Chapter 5, many subsequent results of the sinc basis moment method are presented, most of them generated on the Alliant supercomputer at the University of Illinois Center for Supercomputing Research and Development. In that chapter, many demonstrations of the behavior of this algorithm support the conclusion that, in its present state, successful reconstruction

using the sinc basis method is limited to objects with phase shifts compared with the homogeneous coupling medium of magnitude less than π radians.

1.1 Previous Tomography Studies

1.1a Ultrasonic tomography

Computerized ultrasound tomography began in 1974 with the first attenuation images, reported by Greenleaf et al. (1974). Then followed time of flight refractive index tomograms (Greenleaf et al., 1975). The problem with these images was the assumption that the ultrasonic energy travelled in straight lines, an invalid assumption for ultrasonic propagation in tissue. As a result, diffraction and refraction effects were ignored and in the attenuation images in particular, specular reflection was neglected. This assumption resulted in various artifacts in the reconstructions and low resolution (approximately 20 wavelengths, (Johnson and Tracy, 1983)) because the scattering is not negligible; the scattering correlation distance in tissue is smaller than or of the same order of magnitude as the wavelength. One can not increase the frequency indefinitely to improve resolution because the attenuation is roughly proportional to frequency, so that the penetration depth soon becomes a limiting factor.

One further step was taken by Crawford and Kak (1982), along the ray approach. The received signal, modeled by the sum of a small number of paths, was treated mathematically as the desired ray convolved with a train of impulses. By homomorphic filtering the desired ray was deconvolved from the received signal. Also,

median filtering was used to smooth out spikes in the image thought to be due to refraction. However, the very discussion of rays is inappropriate for describing ultrasound propagation in tissue, as noted above.

In deriving the solution of diffraction problems in a half-space with known boundary conditions on the boundary plane, it was common among optics researchers to expand the field over an angular spectrum of plane waves (Sherman, 1967). Conceptually, the angular spectrum representation of a field is a superposition of plane waves with complex weighting coefficients travelling in all directions, including complex directions (evanescence). The complex weighting coefficients are obtainable exactly as values of the Fourier transform of the field evaluated on a line (or plane) spatially removed from the point of evaluation of the field under consideration. Typically, in diffraction theory this line or plane is called the aperture, where one of the spatial coordinates is fixed at zero and where the field is known along the other(s). However, introduction of a mere propagation factor allows the aperture to be moved to any desired value of the fixed coordinate--for example, the measurement line or plane in a tomography system. The difference between an angular spectrum representation and a straight two- or three-dimensional Fourier transform is the elemental expansion function: for the angular spectrum the function is a plane wave, in particular satisfying $k_x + k_y + k_z = k_0^2$ and usually one (two) of the components of $\vec{k}$ is (are) restricted to be real, while the other may be either real (propagating) or imaginary (evanescent). For the Fourier

transform, the elemental expansion function is a complex exponential with k_x, k_y, and k_z all independent parameters, but all constrained to be real.

Wolf (1969) approximately solved the weak scattering inverse problem by using angular spectra in the following way. Rather than beginning with one of Rayleigh's integral formulas as is common in optical diffraction studies, he began with the Fredholm integral equation for the propagating wave. He expanded not only the field but also the Green function over the angular spectrum. By invoking the Born approximation he made a new identification: the coefficients of the angular spectrum of the scattered field are proportional to values on a semispherical shell of the Fourier transform of the object function. This result has been called the Fourier Diffraction Theorem, and will be treated mathematically in Chapter 2. One could, then, by varying the direction of the incident field (assumed to be a plane wave) fill the entire sphere within the resolution limit and inverse transform the result to obtain the desired object function.

This procedure was soon generalized to include solutions within the Rytov approximation (Iwata and Nagata, 1970). Mueller et al. (1979) offered the first practical computer simulation implementation presentation using the Born approximation; the general procedure was now called "diffraction tomography." The order of number of operations for their technique is $n^2\log(n)$ using FFTs for an nxn reconstruction. The generalization of convolution backprojection to diffraction tomography was given by Devaney (1982). There, either Born or Rytov approximations

in theory could be used, and the computational complexity was $n^3\log(n)$. A further improvement of the Fourier domain reconstruction technique was presented by Nahamoo et al. (1984), where this time the incident field was generalized to an angular spectrum representation; previously, only a single plane wave had been considered. Again the Born approximation was used. The major advantage over the other Fourier domain algorithms was the need for only two object orientations, because each view consisted of an array of scattered field measurements taken for each position of an array of transmitter positions. Computational complexity for this algorithm is n^3. A disadvantage aside from use of the Born approximation is the neglect of information available on the incident side of the object. Indeed, in clinical applications, transmission tomography may be impractical because of the severe attenuation occurring during propagation through the cross section of the body.

1.1b Microwave tomography

Another modality of medical imaging for which diffraction tomography algorithms may apply is microwave imaging. Ghodgaonkar et al. (1983) proposed a moment method solution to the diffraction tomography problem appearing concurrently with the papers of Johnson and Tracy (1983). Ghodgaonkar et al. (1983) used Dirac delta basis functions, while Johnson and Tracy (1983) used sinc basis functions. Both microwave and ultrasonic imaging have the problem that to increase resolution the frequency must be increased, but then the penetration depth decreases because of

attenuation. This problem appears to be more serious for microwave imaging in order to obtain comparable resolution within the same object.

Details of the method of solution for the field used in the tomography algorithm of Ghodgaonkar et al. (1983) were given by Livesay and Chen (1974). It is there that the topic of the singularity of the Green function was raised and had to be dealt with. In this work, as in the present work, the field must be evaluated in the "source" region (i.e., source of the scattered field). The singularity of the Green function was given special treatment; an extra term in the integral equation occurs, deriving from the infinitessimal volume--no matter how small--about the point of field evaluation excluded from the original integration over the object region. This term is due to the singularity of the integrand when the integration parameter coincides with the (internal) point of evaluation of the field. The integrand is singular there because the elements of the Green function dyad have the term

$$\frac{\partial^2}{\partial x_i \partial x_j} \left\{ \frac{e^{jk|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} \right\} \qquad (1.1b.1)$$

which leads to the integral of a multipole term and does not converge because the value of the integral depends upon the shape of the excluded volume referred to above (Van Bladel, 1961). However, if the order of the singularity is less than two it is removable (Wilcox, 1957). In the acoustic case, when density variations are negligible, no term analogous to

$$\frac{\nabla\nabla}{k^2}\left\{\frac{e^{jk|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}\right\} \qquad (1.1b.2)$$

appears because of the differences between the acoustic and electromagnetic equations from which the wave equation is derived (Cavicchi, 1984). Therefore, because only the free-space Green function pertaining to the scalar wave equation is used, the singularity of the Green function is only of order one and consequently is not an issue even when evaluating the field within the object region.

The singularity just discussed is a complicating problem for electromagnetic imaging. Also, one has either to solve for the field by solving a coupled vector wave equation or for both the scalar and the vector potential by solving a scalar and a vector (though uncoupled) wave equation. Attempts to consider the field as scalar (pure polarization) in the inhomogeneous object region are dubious. Slaney et al. (1984) stated that to be able to approximate the <u>field</u> as satisfying a scalar wave equation, the wavelength must be much smaller than the correlation size of the inhomogeneities in the object. Examining the vector wave equation for the electric field

$$\nabla^2\vec{E}(\vec{r}) + k_o^2 n^2 \vec{E}(\vec{r}) - 2\nabla\{\frac{\nabla n}{n} \cdot \vec{E}(\vec{r})\} = 0 \qquad (1.1b.3)$$

the last term must be negligible for a scalar (uncoupled) wave equation to result. Assuming an incident field in, say, the z direction only and the perturbation in the index of

refraction n small, Ishimaru (1978) argued as follows. Because $n = 1 + n_1$,

$$k_o^2 n^2 E_z \simeq k_o^2 E_z + k_o^2 2 n_1 E_z . \qquad (1.1b.4)$$

The order of the last term in Eq. (1.1b.3) can be estimated as follows. Recalling that

$$\nabla(\vec{A}\cdot\vec{B}) = (\vec{B}\cdot\nabla)\vec{A} + (\vec{A}\cdot\nabla)\vec{B} + \vec{B}x(\nabla x\vec{A}) + \vec{A}x(\nabla x\vec{B}) \qquad (1.1b.5)$$

for vectors $\vec{A}$ and $\vec{B}$, if $\nabla n/n$ is used for $\vec{A}$ and $\vec{E}$ for $\vec{B}$, the second term of Eq. (1.1b.5) is easiest to use for estimating the order of the last term in Eq. (1.1b.3): the Laplacian of ∇n in that term is of the order n_1/l_0^2 where l_0 is the correlation distance of n. Because n is of the order one, the last term, $-2\nabla(\nabla n/n\cdot\vec{E})$ is of the order $(2n_1/l_0^2)E_z$. In perturbation problems, for the last term of Eq. (1.1b.3) to be negligible (depolarization effect), a requirement is that $\lambda << l_0$, because $\vec{E}$ approximately satisfies the homogeneous wave equation. That is, if Eq. (1.1b.4) is substituted into Eq. (1.1b.3), the first two terms are both zero order, and the first order, third term, $k_0^2 2 n_1 E_z$, is required to be large compared with the last term, of the order $(2n_1/l_0^2)E_z$; that requirement holds only for $\lambda << l_0$. In that case, the perturbation problem is adequately described by the two zero order terms and the first-order term, or, alternatively, the first two terms of Eq. (1.1b.3). But the condition $\lambda << l_0$ is clearly not satisfied for microwave propagation in tissue. At 3 GHz (for frequencies higher than this attenuation

becomes a major problem) $\lambda = 4.5 \cdot 10^7 / 3 \cdot 10^9 = 1.5$ cm in tissue, while the correlation distances in tissue are of the order of 0.01 cm. Hence, the coupled vector wave equation for the field must be used in electromagnetic tomography. (In (Wilcox, 1957) it is shown that both electric and magnetic fields may be represented by two scalar potentials. But these representations are not valid in the source region (Papas, 1965).)

One other recent microwave tomography study, by Pichot et al. (1985) uses the Fourier Diffraction Theorem described in the previous section but does not invoke the Born approximation. However, what they are imaging is not the object function, but the product of the object function and the field. Furthermore, they present images of an additive mixture of the fields in all directions, a quantity not identifiable with any inherent physical properties of the object.

1.2 Summary

In both ultrasonic and microwave imaging at suitable frequencies, diffraction effects must be accounted for in order to obtain acceptable results. First-order diffraction has been included in several algorithms, nearly all of which incorporate either the Born or the Rytov approximation in order to remain computationally tractible. For theoretical and practical reasons, ultrasound may be better suited than microwaves for use in conjunction with simple tomography algorithms for tissue parameter reconstruction. The sinc basis moment method will be shown to be a way to improve over the Born, and probably the Rytov approxima-

tions within the conditions of its validity. First, however, in the next chapter several conventional tomography algorithms will be examined in detail and compared. This study provides a foundation for understanding the sinc basis method, described in Chapters 3, 4, and 5, and evaluating its performance relative to these first-order diffraction tomography algorithms.

CHAPTER 2

CONVENTIONAL DIFFRACTION TOMOGRAPHY ALGORITHMS

2.0 Introduction

The various tomography algorithms alluded to in Chapter 1 are mathematically related. Although the sinc basis moment method is different from these in that it provides higher-order inverse scattering solutions and in that different computational methods are used, it is important to understand these relations to see where the sinc basis method fits into the body of currently existing tomography algorithms. Furthermore, in Chapter 5 the first iteration reconstruction of the sinc basis moment method is shown computationally to be essentially identical to that of the Born approximation version of the Fourier Diffraction Theorem method; thus, it is worthwhile to examine in detail the conventional diffraction tomography algorithms. The original aspects of this chapter are 1) application of a maximally unified notation to all of the formulations, 2) comments relating the algorithms and other descriptive comments, and 3) the final form of comparison and a comment on the relationship between the Born and Rytov approximations as manifested in the Fourier relation of conventional diffraction tomography.

2.1 The Fourier Projection Theorem

The following discussion begins with a simple statement of the Fourier Projection Theorem valid for straight-path tomography. Beginning with an object inhomogeneity function $\gamma(x,y)$, and its representation in coordinates rotated by angle ϕ with respect

to (x,y) by $\gamma'(u,v)$, a projection of $\gamma(x,y)$ along the direction $\hat{v}$ (see Fig. (2.1.1)) for the two-dimensional case is defined as a line integral of γ' along $\hat{v}$ for a given value of the orthogonal coordinate u:

$$p_v(u) = \int_{-\infty}^{\infty} \gamma'(u,v)dv \tag{2.1.1}$$

where because

$$(x,y) = (u,v)\begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}, \tag{2.1.2}$$

$$\gamma'(u,v) = \gamma(x,y) \tag{2.1.3}$$

and

$$u = x\cos\phi + y\sin\phi \text{ and } v = -x\sin\phi + y\cos\phi. \tag{2.1.4}$$

The Fourier transform of the projection with respect to u, $\tilde{p}_v(k_u)$, where k_u is the transform coordinate corresponding to u

of the spatial domain, can be written

$$\tilde{p}_v(k_u) = \int_{-\infty}^{\infty} p_v(u)e^{-jk_u u}du$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma'(u,v)e^{-j[k_u u+0\cdot v]}\,dudv. \tag{2.1.5}$$

The Fourier Projection Theorem results when the above is recognized as the two-dimensional Fourier transform of the object function evaluated on the k_u axis. Hence,

$$\tilde{p}_v(k_u) = \tilde{\gamma}'(k_u,0) = \tilde{p}_\phi(k_u). \tag{2.1.6}$$

In the last part of the above equation the Fourier transform of

the projection has merely been relabeled with $\tilde{p}_\phi(k_u)$, referencing this projection to the unrotated coordinate system (x,y). Thus, the projection itself is similarly defined with the notation $p_\phi(u)$. The Fourier Projection Theorem can now be written as

$$\tilde{p}_\phi(k_u) = \tilde{\gamma}(k_u\cos\phi,\ k_u\sin\phi). \qquad (2.1.7)$$

Therefore, if at several viewing angles ϕ_i measurements of $p_\phi(u)$ are available for all u, one can obtain the Fourier transform of $\gamma(x,y)$ on polar lines (see Fig. (2.1.2)), where k_x and k_y are the transform variables corresponding to x and y in the spatial domain. If $\tilde{\gamma}$ is known at sufficiently many points, γ is recoverable by inverse Fourier transformation.

2.2 The Fourier Diffraction Theorem

Ultrasonic straight-path tomography is formulated in the following way. If γ is considered to be the index of refraction variation, for example, then $p_\phi(u)/c_0$ is the measured time-of-flight variation for a pulse traveling through the object and detected on the line v = L, where here variation is that with respect to no object present. Or, if γ is considered to be the absorption coefficient of the tissue in the lossless coupling medium, then $p_\phi(u)$ is the measured attenuation of the received pulse (Macovski, 1983). The problem with straight-path tomography is that ultrasonic energy does not travel in straight lines through typical tissue. So in the form given above, the Fourier Projection Theorem does not apply for the case of ultrasonic

tomography. However, first-order diffraction tomography has used a theorem analogous to the Fourier Projection Theorem, modified to account for first-order diffraction effects arising from the interaction of the incident wave with the scattering object represented by γ. Apart from a complex phase shift, knowledge of the Fourier transform of the scattered field measured on the line $v = L$ provides values of the Fourier transform of γ, but on semicircular arcs rather than on radial lines. (The two surfaces, however, merge for the case of vanishing wavelength compared with inhomogeneity correlation distances.) Furthermore, the quantity that γ represents is different from, though easily related to, the index of refraction and absorption (see Section 3.1b):

$$\gamma(x,y) = \omega^2\left(\frac{1}{c^2(x,y)} - \frac{1}{c_0^2}\right) \mp j\omega\frac{\alpha(x,y)}{c(x,y)} \qquad (2.2.1)$$

for $e^{\pm j\omega t}$ time dependence, where $c(x,y)$ and $\alpha(x,y)$ are, respectively, the distributions of speed of sound and absorption in the scattering object, c_0 is the speed of sound in the reference medium, and ω is the radial frequency of the insonifying wave. What has been done for the case of first-order diffraction is a reformulation of the tomography problem in terms of the equation describing the wave nature of ultrasonic propagation in tissue: the scalar, inhomogeneous Helmholtz wave equation. Specifically (see Section 3.1a), if variations in density are ignored,

$$(\nabla^2 + k_0^2)f(x,y) = -\gamma(x,y)f(x,y) \qquad (2.2.2)$$

where f is the total ultrasonic pressure field, γ is as defined

above in Eq. (2.2.1), and $k_0 = \omega/c_0$ is the wavelength in the reference medium. (If density variations are not negligible, the same form of wave equation can be used, but the meanings of the symbols f and γ are changed: $f' = f/\sqrt{\rho}$ where ρ is the local density and

$$\gamma' = \gamma - \sqrt{\rho}\,\nabla^2\left(\frac{1}{\sqrt{\rho}}\right) . \qquad (2.2.3)$$

Here solutions of Eq. (2.2.2) will be considered in an unbounded medium. To this end, the total field is decomposed into the incident field (defined as the total field due to external sources in the absence of a scattering object) and scattered field components:

$$f = f^{inc} + f^{sc} \qquad (2.2.4)$$

where f^{inc} satisfies

$$(\nabla^2 + k_0^2) f^{inc} = 0 \qquad (2.2.5)$$

and therefore f^{sc} satisfies

$$(\nabla^2 + k_0^2) f^{sc} = -\gamma f. \qquad (2.2.6)$$

The solution for the scattered field can be represented as the integral over all induced point sources (the material inhomogeneities can be seen, from Eq. (2.2.6), to act as induced sources of the scattered field) in the scattering region. To see this, denote the field at location $\vec{r}$ due to a point source at $\vec{r}'$ by G. The function G depends only on the distance between $\vec{r}$ and $\vec{r}'$, so it can be written $G(|\vec{r} - \vec{r}'|)$. By definition, G must

satisfy

$$(\nabla^2 + k_0^2)G(|\vec{r}-\vec{r}'|) = -\delta(\vec{r}-\vec{r}'). \tag{2.2.7}$$

Hence, the solution for the scattered field can be written

$$\begin{aligned} f^{sc} &\overset{(a)}{=} f - f^{inc} \\ &\overset{(b)}{=} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f(\vec{r}')G(|\vec{r}-\vec{r}'|)d\vec{r}'. \end{aligned} \tag{2.2.8}$$

This solution satisfies Eq. (2.2.6), as it must, because

$$\begin{aligned} (\nabla^2 + k_0^2)f^{sc} &= (\nabla_{\vec{r}}^2 + k_0^2)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f(\vec{r}')G(|\vec{r}-\vec{r}'|)d\vec{r}' \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f(\vec{r}')(\nabla_{\vec{r}}^2 + k_0^2)G(|\vec{r}-\vec{r}'|)d\vec{r}' \\ &= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f(\vec{r}')\delta(\vec{r}-\vec{r}')d\vec{r}' \\ &= -\gamma(\vec{r})f(\vec{r}). \end{aligned} \tag{2.2.9}$$

It is well known that in three dimensions

$$G(|\vec{r}-\vec{r}'|) = \frac{e^{\mp jk_0|\vec{r}-\vec{r}'|}}{4\pi|\vec{r}-\vec{r}'|} \tag{2.2.10}$$

and in two dimensions (see Section 3.1d)

$$G(|\vec{r}-\vec{r}'|) = \mp\frac{j}{4}H_0^{(\frac{2}{1})}(k_0|\vec{r}-\vec{r}'|) \tag{2.2.11}$$

for $e^{\pm j\omega t}$ time dependence assumed.

The focus of this study is on two dimensions; all of the theory can simply be extended to three dimensions if desired. Also, it will be assumed that the time dependence is $e^{-j\omega t}$, as

opposed to (Johnson and Tracy, 1983) and Chapter 3, but in agreement with other investigators such as Devaney (1982) and Pan and Kak (1983). The reason for doing this is because of the (arbitrary) definition of the sign of $\sqrt{-1}$ in standard Fourier transform definitions. The final results would be the same, had $e^{+j\omega t}$ been assumed, but then inverse FFTs of scattered fields would have to be taken in the final equations.

The solution of the homogeneous wave equation

$$(\nabla^2 + k_0^2) f^{inc} = 0 \qquad (2.2.12)$$

is a superposition of plane waves, or an angular spectrum. For a plane wave propagating in direction $\vec{k}$, f^{inc} has the form

$$f^{inc} = f_0(\omega) e^{-j\vec{k}\cdot\vec{r}} \qquad \text{where} \begin{cases} \vec{k} = (k_x, k_y) \\ k_x^2 + k_y^2 = k_0^2 \\ \vec{r} = (x,y) \ . \end{cases} \qquad (2.2.13)$$

To obtain the desired relation between the Fourier transform of the scattered field and that of the object function γ, Wolf (1969) expanded the Green function over the angular spectrum. In two dimensions this expansion is

$$\frac{j}{4} H_0^{(1)}(k_0|\vec{r}-\vec{r}'|) = \frac{j}{4\pi} \int_{-\infty}^{\infty} \frac{e^{j[k_x(x-x') + k_y^*|y-y'|]}}{k_y^*} \, dk_x$$

$$\text{where } k_y^* = \sqrt{k_0^2 - k_x^2}. \qquad (2.2.14)$$

Substituting this expansion into Eq. (2.2.8) gives

$$f^{sc}(x,y) = \frac{j}{4\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(x',y')f(x',y')$$

$$\cdot\int_{-\infty}^{\infty} \frac{e^{j[k_x(x-x') + k_y^*|y-y'|]}}{k_y^*} \, dk_x dx' dy' \qquad (2.2.15)$$

Up through Eq. (2.2.15), no first-order approximations have been made in the solution of Eq. (2.2.6). However, in order to obtain a physically meaningful quantity from the process of reconstruction, such an approximation is necessary in order to proceed further; that is, γ needs to be able to be separated out from the product γf. Two methods of linearization (note the product of two unknowns under the integral) have been proposed.

2.2a The Born approximation

The first Born approximation replaces the total field in Eq. (2.2.15) by the incident field. In Eq. (2.2.8) this implies that

$$\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f^{sc}(\vec{r}')G(|\vec{r}-\vec{r}'|)d\vec{r}'\right| \ll$$

$$\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f^{inc}(\vec{r}')G(|\vec{r}-\vec{r}'|)d\vec{r}'\right|. \qquad (2.2a.1)$$

The incident field can be substituted as an approximation of the total field under the integral sign in Eq. (2.2.8), and the result of integration identified as the scattered field, to obtain the first Born approximation. Thus,

$$f^{sc} \cong f_*^{sc} \triangleq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f^{inc}(\vec{r}')G(|\vec{r}-\vec{r}'|)d\vec{r}' \qquad (2.2a.2)$$

in two dimensions. The symbol f_*^{sc} represents the right-hand side, that is, the result of substituting the incident field for the total field under the integral sign. The Born approximation

results when it is assumed that $f^{sc} = f_*^{sc}$. For the case of a single plane wave of amplitude $f_0(\omega)$ in two dimensions,

$$f_*^{sc} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f_0(\omega)e^{j\vec{k}\cdot\vec{r}'} G(|\vec{r}-\vec{r}'|)d\vec{r}'. \qquad (2.2a.3)$$

(The designation of dependence of the plane wave amplitude on ω is simply a reminder that a wave of any well-behaved time dependence can be decomposed into its temporal frequency components.)

2.2b The Rytov approximation

The assumption for the Rytov approximation is made while still working with the differential form, Eq. (2.2.2). Note that γ can be expressed as

$$\gamma(\vec{r}) = k_0^2[n^2(\vec{r})-1] \qquad (2.2b.1)$$

where $n(\vec{r})$ is the index of refraction $c_0/c(\vec{r})$. (Here, for lossy objects, $n(\vec{r})$ and $c(\vec{r})$ will be complex.) Now temporarily assume that the only restriction upon f^{inc} is that it satisfies $(\nabla^2 + k_0^2)f^{inc} = 0$ (Eq. (2.2.5)). Express the incident and total fields as

$$f^{inc}(\vec{r}) = f_0(\omega)e^{j[\psi_0(\vec{r})-\omega t]} \qquad (2.2b.2)$$

$$f(\vec{r}) = f_0(\omega)e^{j[\psi(\vec{r})-\omega t]} \qquad (2.2b.3)$$

(where, of course, in general $\psi(\vec{r})$ must be complex). Substitute Eq. (2.2b.3) into Eq. (2.2.2):

$$(\nabla^2+k_0^2)f(\vec{r}) = (\nabla^2+k_0^2)f_0(\omega)e^{j[\psi(\vec{r})-\omega t]} = -\gamma(\vec{r})f_0(\omega)e^{j[\psi(\vec{r})-\omega t]}. \qquad (2.2b.4)$$

Cancelling the common factor $f_0(\omega)e^{-j\omega t}$,

$$\nabla^2 e^{j\psi(\vec{r})} = -[k_0^2 + \gamma(\vec{r})]e^{j\psi(\vec{r})}. \qquad (2.2b.5)$$

Expanding the Laplacian,

$$\nabla^2 e^{j\psi(\vec{r})} = \nabla\{j\nabla\psi(\vec{r})e^{j\psi(\vec{r})}\} = \{j\nabla^2\psi(\vec{r}) - [\nabla\psi(\vec{r})]^2\}e^{j\psi(\vec{r})} \quad (2.2b.6)$$

so that Eq. (2.2b.5) becomes

$$-[\nabla\psi(\vec{r})]^2 + j\nabla^2\psi(\vec{r}) = -[k_0^2 + \gamma(\vec{r})]. \quad (2.2b.7)$$

Substituting Eq. (2.2b.2) into Eq. (2.2.5),

$$(\nabla^2+k_0^2) f^{inc}(\vec{r}) = f_0(\omega)e^{-j\omega t}\{\nabla^2 e^{j\psi_0(\vec{r})} + k_0^2 e^{j\psi_0(\vec{r})}\} = 0. \quad (2.2b.8)$$

Rearranging and expanding $\nabla^2 e^{j\psi_0(\vec{r})}$,

$$\nabla^2 e^{j\psi_0(\vec{r})} = \nabla\{j\nabla\psi_0(\vec{r})e^{j\psi_0(\vec{r})}\} = -k_0^2 e^{j\psi_0(\vec{r})} \quad (2.2b.9)$$

or

$$j\nabla^2\psi_0(\vec{r}) - [\nabla\psi_0(\vec{r})]^2 = -k_0^2 \quad (a) \quad (2.2b.10)$$

$$[\nabla\psi_0(\vec{r})]^2 - j\nabla^2\psi_0(\vec{r}) = k_0^2. \quad (b) \quad (2.2b.10)$$

Adding Eq. (2.2b.10b) to Eq. (2.2b.7),

$$-[\nabla\psi(\vec{r})]^2 + [\nabla\psi_0(\vec{r})]^2 + j\nabla^2[\psi(\vec{r}) - \psi_0(\vec{r})] = -\gamma(\vec{r}). \quad (2.2b.11)$$

Define the phase fluctuation

$$\psi'(\vec{r}) = \psi(\vec{r}) - \psi_0(\vec{r}). \quad (2.2b.12)$$

Then

$$[\nabla\psi'(\vec{r})]^2 = [\nabla\psi(\vec{r})]^2 - 2\nabla\psi(\vec{r}) \cdot \nabla\psi_0(\vec{r}) + [\nabla\psi_0(\vec{r})]^2 =$$

$$[\nabla\psi(\vec{r})]^2 - 2\nabla\psi'(\vec{r}) \cdot \nabla\psi_0(\vec{r}) - [\nabla\psi_0(\vec{r})]^2 \quad (2.2b.13)$$

so that

$$-[\nabla\psi(\vec{r})]^2 + [\nabla\psi_0(\vec{r})]^2 = -[\nabla\psi'(\vec{r})]^2 - 2\nabla\psi'(\vec{r}) \cdot \nabla\psi_0(\vec{r}), \quad (2.2b.14)$$

Substituting Eq. (2.2b.14) into Eq. (2.2b.11),

$$-2\nabla\psi_0(\vec{r}) \cdot \nabla\psi'(\vec{r}) + j\nabla^2\psi'(\vec{r}) = -\gamma(\vec{r}) + [\nabla\psi'(\vec{r})]^2. \quad (2.2b.15)$$

At this stage the Rytov approximation is made: if $|\nabla\psi'(\vec{r})|/k_0$ is of the same order as the fluctuation in the index of refraction,

$n_1(\vec{r})$, where $n(\vec{r}) = 1 + n_1(\vec{r}) = c_0/c(\vec{r})$ so that $n_1(\vec{r}) = (c_0 - c(\vec{r}))/c(\vec{r})$ then $[\nabla\psi'(\vec{r})]^2$ is a second-order quantity in $n_1(\vec{r})$ and can be ignored. Note that

$$\begin{aligned}\gamma(\vec{r}) &= k_0^2[n^2(\vec{r})-1] \\ &= k_0^2\{[1+n_1(\vec{r})]^2 - 1\} \\ &= k_0^2[2n_1(\vec{r}) + n_1^2(\vec{r})]\end{aligned} \qquad (2.2b.16)$$

contains the <u>first</u>-order term $2k_0^2 n_1(\vec{r})$. Thus, because

$$|\nabla\psi'(\vec{r})|/k_0 \sim n_1(\vec{r}) << 1 \qquad (2.2b.17)$$

(a weak scattering assumption) the condition for validity of the Rytov approximation is

$$|\nabla\psi'(\vec{r})|/k_0 << 1. \qquad (2.2b.18)$$

For an increment $\Delta\psi'(\vec{r})$ in $\psi'(\vec{r})$ over a distance Δx, the condition reads

$$\lambda\Delta\psi'(\vec{r})/\Delta x << 2\pi. \qquad (2.2b.19)$$

Thus, if Δx is chosen to be the order of a wavelength, then $\psi'(\vec{r})$ can not in that interval change much compared with 2π. Thus, it is the phase change <u>per wavelength</u> that must be small; no condition is placed upon the <u>total</u> (integrated) phase and amplitude changes across the scatterer, as was done in the Born approximation (see Eq. (2.2a.1)). There the phase condition translates to a maximum phase difference of π across the object (Slaney and Kak, 1985). Proceeding with the simplified Eq. (2.2b.15),

$$-2\nabla\psi_0(\vec{r}) \cdot \nabla\psi'(\vec{r}) + j\nabla^2\psi'(\vec{r}) \simeq -\gamma(\vec{r}). \qquad (2.2b.20)$$

Now assume f^{inc} to be a plane wave. It is questionable at this stage what results could be obtained without this assumption. That is, it is doubtful whether without assuming an incident plane wave the differential equation for W (see below) would be simple enough to solve and obtain the simple relation to the Born solution, as shown below. In the plane wave case,

$$\psi_0(\vec{r}) = \vec{k}\cdot\vec{r} \rightarrow \nabla\psi_0(\vec{r}) = \vec{k}. \tag{2.2b.21}$$

Eq. (2.2b.20) now becomes

$$-2\vec{k} \cdot \nabla\psi'(\vec{r}) + j\nabla^2\psi'(\vec{r}) = -\gamma(\vec{r}). \tag{2.2b.22}$$

Define W to be the function satisfying

$$\psi'(\vec{r}) = e^{-j\vec{k}\cdot\vec{r}} \cdot W(\vec{r}). \tag{2.2b.23}$$

Then

$$\nabla\psi'(\vec{r}) = -j\vec{k}e^{-j\vec{k}\cdot\vec{r}}W(\vec{r}) + e^{-j\vec{k}\cdot\vec{r}}\nabla W(\vec{r}) = [-j\vec{k}W(\vec{r}) + \nabla W(\vec{r})]e^{-j\vec{k}\cdot\vec{r}} \tag{2.2b.24}$$

and

$$\nabla^2\psi'(\vec{r}) = \{-j\vec{k}\cdot\nabla W(\vec{r}) + \nabla^2 W(\vec{r}) - j\vec{k} \cdot [-j\vec{k}W(\vec{r}) + \nabla W(\vec{r})]\}e^{-j\vec{k}\cdot\vec{r}}$$
$$= \{-j\vec{k}\cdot\nabla W(\vec{r}) - k_0^2 W(\vec{r}) - j\vec{k}\cdot\nabla W(\vec{r}) + \nabla^2 W(\vec{r})\}e^{-j\vec{k}\cdot\vec{r}} \tag{2.2b.25}$$

Eq. (2.2b.24) yields

$$\vec{k}\cdot\nabla\psi'(\vec{r}) = [-jk_0^2 W(\vec{r}) + \vec{k}\cdot\nabla W(\vec{r})]e^{-j\vec{k}\cdot\vec{r}} \tag{2.2b.26}$$

Substituting Eq. (2.2b.25) and Eq. (2.2b.26) into Eq. (2.2b.22) yields

$$-2\vec{k}\cdot\nabla\psi'(\vec{r}) + j\nabla^2\psi'(\vec{r}) =$$
$$= \{2jk_0^2 W(\vec{r}) - 2\vec{k}\cdot\nabla W(\vec{r}) + j\nabla^2 W(\vec{r}) - jk_0^2 W(\vec{r}) + 2\vec{k}\cdot\nabla W(\vec{r})\}e^{-j\vec{k}\cdot\vec{r}}$$
$$= -\gamma(\vec{r}) \tag{2.2b.27}$$

or

$$(\nabla^2 + k_0^2)W(\vec{r}) = j\gamma(\vec{r})e^{j\vec{k}\cdot\vec{r}} \tag{2.2b.28}$$

The solution for W is obtained by the same method that f^{sc} in Eq.

(2.2.8) was obtained from Eq. (2.2.2). In two dimensions,

$$W(\vec{r}) = j\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')e^{j\vec{k}\cdot\vec{r}'}G(|\vec{r}-\vec{r}'|)d\vec{r}'. \qquad (2.2b.29)$$

Substituting Eq. (2.2b.29) into Eq. (2.2b.23),

$$\psi'(\vec{r}) = je^{-j\vec{k}\cdot\vec{r}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')e^{j\vec{k}\cdot\vec{r}'}G(|\vec{r}-\vec{r}'|)d\vec{r}'$$

$$= \frac{j}{f_0(\omega)e^{j\vec{k}\cdot\vec{r}}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(\vec{r}')f_0(\omega)e^{j\vec{k}\cdot\vec{r}'}G(|\vec{r}-\vec{r}'|)d\vec{r}'. \qquad (2.2b.30)$$

Noting the resemblance to Eq. (2.2a.3), the phase perturbation is seen to be

$$j\psi'(\vec{r}) \simeq \frac{f_*^{sc}(\vec{r})}{f^{inc}(\vec{r})}. \qquad (2.2b.31)$$

2.2c Completion of derivation of the Fourier Diffraction Theorem

Thus, in terms of reconstruction algorithms it is sufficient to consider the solution of Eq. (2.2a.3). The crucial difference is which quantity is used (obtained by measurement) for input to the algorithm. For the Born approximation, the measured complex scattered field $f^{sc}(\vec{r})$ as defined in Eq. (2.2.8a) is used. For the Rytov approximation it is necessary to have the complex phase of the total field minus that of the incident field: this is obtained by taking the difference in the complex logs:

$$j\psi'(\vec{r}) = \ln f(\vec{r}) - \ln f^{inc}(\vec{r})$$

$$= \ln\left\{\frac{f^{sc}(\vec{r})}{f^{inc}(\vec{r})} + 1\right\} \qquad (2.2c.1)$$

and $$f_*^{sc}(\vec{r}) \simeq jf^{inc}(\vec{r})\psi'(\vec{r}). \qquad (2.2c.2)$$

If $|f^{sc}|$ is of the order of or larger than $|f^{inc}|$, problems with phase wrapping will result (Slaney and Kak, 1985). Note that, for small $|f^{sc}|$ compared with $|f^{inc}|$, Eqs. (2.2b.31) and (2.2c.1) are in agreement despite their difference in form:

$$j\psi'(\vec{r}) = \ln\{\frac{f^{sc}(\vec{r})}{f^{inc}(\vec{r})} + 1\} \simeq \frac{f_*^{sc}(\vec{r})}{f^{inc}(\vec{r})} \rightarrow e^{\frac{f_*^{sc}(\vec{r})}{f^{inc}(\vec{r})}} \simeq 1 + \frac{f_*^{sc}(\vec{r})}{f^{inc}(\vec{r})}. \tag{2.2c.3}$$

This is only the series expansion for the exponential with small argument. Apparently, the difference between the Born and Rytov approximations can be thought of in terms of which of the two quantities better represents the right-hand side of Eq. (2.2a.3) (f_*^{sc}):

$$f - f^{inc} \text{ (BORN) or } f^{inc}\ln\{\frac{f^{sc}}{f^{inc}} + 1\} \text{ (RYTOV)} \tag{2.2c.4}$$

(Slaney and Kak, 1985). Alternatively, noting that for $|x| < 1$, $\ln(1+x) = x - x^2/2 + \ldots$; evidently the first Born approximation, in asserting that $f_*^{sc} = f^{sc}$, keeps only the first term of the Taylor expansion of the definition of f_*^{sc} under the Rytov approximation (Eq. 2.2c.4). Thus, they are approximately equivalent for weak scattering ($|f^{sc}| << |f^{inc}|$). Using the incident field for f under the integral sign in Eq. (2.2.15), and choosing it to be a single plane wave propagating in direction $\vec{k}$ gives

$$f_{*\phi}^{sc}(x,y) = \frac{jf_0^{(\omega)}}{4\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma(x',y')e^{j\vec{k}\cdot\vec{r}'} \cdot \int_{-\infty}^{\infty} \frac{e^{j[k_x(x-x') + k_y^*|y-y'|]}}{k_y^*} dk_x dx' dy' \tag{2.2c.5}$$

where

$$k_y^* = \sqrt{k_0^2 - k_x^2}. \qquad (2.2c.6)$$

The subscript ϕ on $f_{*\phi}^{sc}$ indicates the dependence of the scattered field on the view direction ϕ of the incident field of the u axis with respect to the x axis. If the scattered field is measured on a line perpendicular to the direction of the incident field, and a distance L from the origin, f^{sc} is being measured on the rotated coordinates (u,v = L) (see Fig. (2.1.1)). Hence, $\vec{k}\cdot\vec{r}' = k_0 v'$ in Eq. (2.2c.5). As opposed to previous notation, $f_{*\phi}^{sc}$ will not be primed when the point of evaluation is indicated with respect to (u,v) coordinates, simply because this is the only system of specification that will be used for specifying the scattered field, as opposed to, for example, γ in Eq. (2.1.3).

Consider the angular spectrum integral representation of the Hankel function in Eq. (2.2c.5). First note that

$$H_0^{(1)}[k_0\sqrt{(x-x')^2 + (y-y')^2}] = H_0^{(1)}[k_0\sqrt{(u-u')^2 + (v-v')^2}] \qquad (2.2c.7)$$

where (x,y) <=> (u,v) and (x',y') <=> (u',v') are the coordinates for the points $\vec{r}$ and $\vec{r}'$, respectively, in the fixed and rotated coordinate systems. The point is, of course, that one can write the integral form in Eq. (2.2.14) as

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{dk_x}{k_y^*} e^{j[k_x(x-x') + k_y^*|y-y'|]} =$$

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{dk_u}{k_v^*} e^{j[k_u(u-u') + k_v^*|v-v'|]} \qquad (2.2c.8)$$

and again the restriction $k_v^* = \sqrt{k_0^2 - k_u^2}$.

Thus, one can write

$$f_{*\phi}^{sc}(u,L) = \frac{jf_0(\omega)}{4\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx'dy'\gamma(x',y')e^{jk_0v'} \cdot \int_{-\infty}^{\infty} \frac{dk_u}{k_v^*} e^{j[k_u(u-u') + k_v^*|L-v'|]} . \quad (2.2c.9)$$

The distance L is chosen greater than all v within the object region. Therefore, $|L - v'| = L - v'$. Rearranging,

$$f_{*\phi}^{sc}(u,L) = \frac{jf_0(\omega)}{4\pi}\int_{-\infty}^{\infty} \frac{dk_u}{k_v^*} e^{j[k_uu + k_v^*L]} \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx'dy'\gamma(x',y')e^{j[-k_uu' + (k_0-k_v^*)v']} . \quad (2.2c.10)$$

Let

$$\hat{s} = \left(\frac{k_u\hat{u} + k_v^*\hat{v}}{k_0}\right) \text{ and } \hat{s}_0 = \hat{v} . \quad (2.2c.11)$$

Then

$$e^{-j[k_uu' - (k_0-k_v^*)v']} = e^{-jk_0(\hat{s}-\hat{s}_0)\cdot\vec{r}'} = e^{-j(k_xx' + k_yy')} \quad (2.2c.12)$$

where

$$(k_x,k_y) = k_0(\hat{s}-\hat{s}_0) . \quad (2.2c.13)$$

If k_x and k_y are so defined, then

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx'dy'\gamma(x',y') e^{-j[k_uu' - (k_0-k_v^*)v']} = \tilde{\gamma}(k_x,k_y) = \tilde{\gamma}'(k_u,k_v^*) . \quad (2.2c.14)$$

Hence,

$$f_{*\phi}^{sc}(u,L) = \frac{jf_0(\omega)}{4\pi}\int_{-\infty}^{\infty}\frac{dk_u}{k_v^*}e^{j[k_u u + k_v^* L]} \cdot \tilde{\gamma}[k_0(\hat{s}-\hat{s}_0)]. \quad (2.2c.15)$$

Now take the one-dimensional Fourier transform with respect to u: (recognize the inverse Fourier transform on the right-hand side)

$$\tilde{f}_{*\phi}^{sc}(k_u,L) = \int_{-\infty}^{\infty} f_{*\phi}^{sc}(u,L)e^{-jk_u u}du$$

$$= \frac{jf_0(\omega)}{2} \cdot \frac{e^{jk_v^* L}}{k_v^*} \cdot \tilde{\gamma}[k_0(\hat{s}-\hat{s}_0)] . \quad (2.2c.16)$$

Rearranging,

$$\frac{2k_v^* e^{-jk_v^* L}}{jf_0(\omega)}\int_{-\infty}^{\infty} f_{*\phi}^{sc}(u,L)e^{-jk_u u}du = \tilde{\gamma}[k_0(\hat{s}-\hat{s}_0)]$$

$$\int_{-\infty}^{\infty}\frac{2k_v^* e^{-jk_v^* L}}{jf_0(\omega)} f_{*\phi}^{sc}(u,L)e^{-jk_u u}du. = \quad (2.2c.17)$$

In the last step the term $2k_v^* e^{-jk_v^* L}/(jf_\phi(\omega))$ could be moved inside the integral sign even though it depends on k_u (via k_v^*) because it does <u>not</u> depend on u, the integration variable. Define

$$D_\phi(u,\omega,k_u) = \frac{2k_v^* e^{-jk_v^* L}}{jf_0(\omega)} f_{*\phi}^{sc}(u,L). \quad (2.2c.18)$$

Then

$$\tilde{D}_\phi(k_u,\omega) \triangleq \int_{-\infty}^{\infty} D_\phi(u,\omega,k_u)e^{-jk_u u}du = \tilde{\gamma}[k_0(\hat{s}-\hat{s}_0)]. \quad (2.2c.19)$$

Note that $\tilde{D}_\phi(k_u,\omega)$ is the product of a Fourier domain (k_u) function and a spatial domain (u) function. The function $\tilde{D}_\phi(k_u,\omega)$ transforms out the u dependence, to yield a function of only k_u (and ω). Note that in this usage (Eqs. (2.2c.17) and

(2.2c.18)) this is perfectly legal. For the Born and Rytov approximations, $f_{*\phi}^{sc}(u,L)$ has the following meanings:

$$f_{*\phi}^{sc}(u,L) = \begin{cases} f_\phi(u,L) - f_\phi^{inc}(u,L) = f_\phi^{sc}(u,L) & \text{Born approx.} \\ f_\phi^{inc}(u,L)\ \ln\{\dfrac{f_\phi^{sc}(u,L)}{f_\phi^{inc}(u,L)} + 1\} & \text{Rytov approx.} \end{cases} \tag{2.2c.20}$$

Again, the $*\phi$ simply indicates that $f_{*\phi}^{sc}(u,L)$ is only a <u>symbol representation</u> for the right-hand side of Eq. (2.2c.5) in the coordinate system and for the incident field defined by the rotation angle ϕ. If it is assumed that $f_\phi^{sc} = f_{*\phi}^{sc}$ then the Born approximation has been made. Or, if one assumes that

$$f_\phi^{inc}\ln\{\frac{f_\phi^{sc}}{f_\phi^{inc}} + 1\} = f_{*\phi}^{sc} \tag{2.2c.21}$$

the Rytov approximation has been made. Noting Eq. (2.2c.3a), Eqs. (2.2c.18) and (2.2c.20) can be rewritten similarly to (Devaney, 1982):

$$D_\phi(u,\omega,k_u) = -j2k_v^* e^{-j(k_v^* - k_0)L}\,\Gamma_\phi(u,\omega) \tag{2.2c.22}$$

where

$$\Gamma_\phi(u,\omega) = \begin{cases} \dfrac{f_\phi^{sc}(u,L)}{f_0(\omega)e^{jk_0L}} & \text{Born approx.} \\ j\psi'_\phi(u,L) & \text{Rytov approx.} \end{cases} \tag{2.2c.23}$$

Note that Γ_ϕ can be rewritten as

$$\Gamma_\phi \overset{(a)}{=} \begin{cases} \dfrac{f-f^{inc}}{f^{inc}} & (B) \\ j\psi' & (R) \end{cases} \quad \overset{(b)}{=} \begin{cases} \dfrac{e^{j\psi}}{e^{j\psi_0}} - 1 & (B) \\ j\psi' & (R) \end{cases} \qquad (2.2c.24)$$

$$\overset{(c)}{=} \begin{cases} e^{j\psi'} - 1 & (B) \\ j\psi' & (R) \end{cases}$$

which are obviously approximately equivalent if $|j\psi'| << 1$; they differ only in the higher order Taylor expansion terms of $e^{j\psi'}$. For later reference, and from Eqs. (2.2c.19) and (2.2c.22), note that

$$\tilde{D}_\phi(k_u,\omega) = -j2k_v^* e^{-j(k_v^*-k_0)L}\,\tilde{\Gamma}_\phi(k_u,\omega) \qquad (2.2c.25)$$

where

$$\tilde{\Gamma}_\phi(k_u,\omega) = \int_{-\infty}^{\infty} \Gamma_\phi(u,\omega) e^{-jk_u u}\,du. \qquad (2.2c.26)$$

From Eqs. (2.2c.18), (2.2c.19), and (2.2c.20) it is observed that information is available about the Fourier transform of the object inhomogeneity function from the measured data of the scattered field either through the Born or through the Rytov approximation. But notice from the definition of k_v^* that if $|k_u| > k_0$, k_v^* becomes purely imaginary. From Eq. (2.2c.14) it is clear that this would mean the Laplace rather than the Fourier transform of γ is being evaluated, for which there are no available methods of numerical computation. For this reason the restriction $|k_u| \leq k_0$ is made, thereby ignoring waves evanescent in the $\hat{v}$ direction, in order to guarantee a Fourier transform relationship. In the

following, keep in mind Eq. (2.2c.11). Note that the Fourier transform in rotated coordinates of a function is equal to the Fourier transform in unrotated coordinates of the function, evaluated at the same location in the frequency plane, expressed in terms of the unrotated coordinates. Also, if it were not for the complexity of notation, each u, k_u, and k_v^* should actually be given a ϕ subscript, because each insonification angle produces a different definition of (u,v) and consequently (k_u,k_v^*) with respect to the (k_x,k_y) (unrotated) coordinate system. Also, throughout this study, the ϕ subscripting of a function indicates that it (for example, an ultrasonic field distribution) depends on the viewing angle (incident field generated). But notice that, as γ is an intrinsic material property of the inhomogeneous medium, it does not depend on the viewing angle, and so is not subscripted with ϕ.

Values of $\tilde{\gamma}$ are available on the following semicircles defined by $\tilde{\gamma}(k_u, -(k_0 - k_v^*))$ where $k_v^* = \sqrt{k_0^2 - k_u^2}$ and $|k_u| \leq k_0$ (see Fig. (2.2c.1)). Noting that (u,v) <=> (k_u,k_v),

$$k_v = -k_0 + k_v^* \qquad (2.2c.27a)$$

or

$$(k_v + k_0)^2 + k_u^2 = k_0^2. \qquad (2.2c.27b)$$

The object function $\tilde{\gamma}$ is known a semicircle centered at $(k_u = 0, k_v = -k_0)$. Because above the positive square root was taken for k_v^*, $(k_v + k_0)$ must be constrained to be positive, or $k_v > -k_0$; this is the reason for obtaining only the half of the k_0-circle in Fig. (2.2c.1) containing the origin. The Fourier space can be filled up within a $\sqrt{2}\,k_0$ circle by taking scattered field

measurements for several different viewing angles ϕ (see Fig. (2.2c.2)). The Fourier space of γ can also be filled up by varying the frequency (and therefore k_0) at a particular viewing angle (see Fig. (2.2c.3)). Either method simultaneously reduces the problem of nonuniqueness in source determination inherent in a monoview scattered field measurement inverse scattering system. Note: it can be shown that in reflection mode tomography, the outer (higher spatial freqfuency) semicircle for frequencies between $\sqrt{2}\,k_0$ and $2k_0$ is obtained (Roberts and Kak, 1985) (see Fig. (2.2c.4)).

2.3 The Filtered Backprojection Algorithm

It is obvious that in using the Fourier Diffraction Theorem one is faced with spatial frequency domain interpolation problems (though from different contours), just as in the case of use of the Fourier Projection Theorem in straight-path tomography, where the contours are radial lines rather than offset semicircles. In conventional tomography, this problem of interpolation has been avoided, with superior results, by using the convolution (filtered) back-projection algorithm. Beginning with Eq. (2.1.6),

$$\tilde{p}_\phi(k_u) = \tilde{\gamma}'(k_u,0) = \tilde{\gamma}(k_u,\phi) \qquad (2.3.1)$$

where the last change of notation shows the important step of realizing that $\tilde{\gamma}(k_x = k_u\cos\phi,\ k_y = k_u\sin\phi) = \tilde{\gamma}(k_u,\phi)$ is a polar coordinates evaluation of $\tilde{\gamma}$. Thus, in the unrotated coordinate system k_u functions as the radial spatial frequency coordinate associated with view angle ϕ. Hence, its inverse transform can be written as

$$\gamma(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{\gamma}(k_x,k_y)e^{j(k_x x + k_y y)}\,dk_x dk$$

$$= \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{\infty} \tilde{\gamma}(k_u,\phi)e^{jk_u u}k_u dk_u d\phi$$

$$= \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{\infty} \tilde{\gamma}(k_u,\phi)e^{jk_u(x\cos\phi + y\sin\phi)}k_u dk_u d\phi$$

$$= \frac{1}{(2\pi)^2}\int_0^{\pi}\int_0^{\infty} \tilde{\gamma}(k_u,\phi)e^{jk_u(x\cos\phi + y\sin\phi)}k_u dk_u d\phi \; +$$

$$\frac{1}{(2\pi)^2}\int_0^{\pi}\int_0^{\infty} \tilde{\gamma}(k_u,\phi+180^o)e^{\overbrace{jk_u[x\cos(\phi+180^o) + y\sin(\phi+180^o)]}^{-jk_u(x\cos\phi + y\sin\phi) = -jk_u u.}}$$

(2.3.2)

Note that

$$\tilde{\gamma}(k_u,\phi+180^o) = \tilde{\gamma}(k_u,-\phi) = \tilde{\gamma}(-k_u,\phi) \qquad (2.3.3)$$

so therefore k_u can be run from $-\infty$ to $+\infty$ by replacing $k_u dk_u$ by $|k_u|dk_u$. Thus,

$$\gamma(x,y) = \frac{1}{(2\pi)^2}\int_0^{\pi}\int_{-\infty}^{\infty} \tilde{\gamma}(k_u,\phi)\,|k_u|\,e^{jk_u u}\,dk_u d\phi. \qquad (2.3.4)$$

By Eq. (2.3.1),

$$\gamma(x,y) = \frac{1}{2\pi}\int_0^{\pi}\{\frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{p}_\phi(k_u)\,|k_u|\,e^{jk_u u}dk_u\}d\phi$$

$$= \frac{1}{2\pi}\int_0^{\pi} Q_\phi(u)d\phi \qquad (2.3.5)$$

where

$$Q_\phi(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{p}_\phi(k_u)\,|k_u|\,e^{jk_u u}\,dk_u \qquad (2.3.6)$$

or,

$$\gamma(x,y) = \frac{1}{2\pi}\int_0^{\pi} Q_\phi(x\cos\phi + y\sin\phi)d\phi. \qquad (2.3.7)$$

In order to proceed further it is noted that in any practical application $Q_\phi(u)$ can not be computed using <u>all</u> spatial frequencies k_u. It is assumed that γ is a bandlimited function: $\tilde{\gamma}(k_u,\phi) = 0$ for $|k_u| > \Omega_{max_\gamma}$. Then $\tilde{p}_\phi(k_u)$ also is zero for $|k_u| > \Omega_{max_\gamma}$. Therefore,

$$Q_\phi(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{p}_\phi(k_u)\tilde{h}_{\Omega_{max_\gamma}}(k_u)e^{jk_u u}dk_u \qquad (2.3.8)$$

where

$$\tilde{h}_\Omega(k) = |k|\tilde{b}_\Omega(k) \qquad (2.3.9)$$

where

$$\tilde{b}_\Omega(k) = \begin{cases} 1 & |k| < \Omega \\ 0 & \text{otherwise.} \end{cases} \qquad (2.3.10)$$

Let us determine h(u):

$$h_{\Omega_{max_\gamma}}(u) = \frac{1}{2\pi}\left\{-\int_{\Omega_{max_\gamma}}^{0} k_u e^{jk_u u}dk_u + \int_0^{\Omega_{max_\gamma}} k_u e^{jk_u u}dk_u\right\}$$

$$\mu = k_u \qquad dv = e^{jk_u u}dk_u$$

$$d\mu = dk_u \qquad v = \frac{1}{ju}e^{jk_u u} \qquad (2.3.11)$$

Integrating by parts,

$$h_{\Omega_{max_\gamma}}(u) = \frac{1}{2\pi}\left\{ - 0 - \frac{(-\Omega_{max_\gamma}e^{-j\Omega_{max_\gamma}u})}{ju} - \frac{1}{ju}\int_{-\Omega_{max_\gamma}}^{0} e^{jk_u u}dk_u \right.$$

$$\left. + \frac{\Omega_{max_\gamma}}{ju}e^{j\Omega_{max_\gamma}u} - \frac{1}{ju}\int_0^{\Omega_{max_\gamma}} e^{jk_u u}dk_u \right\} \qquad (2.3.12a)$$

$$h_{\Omega_{max_\gamma}}(u) = \frac{1}{2\pi}\left\{ \frac{-\Omega_{max_\gamma}}{ju} e^{-j\Omega_{max_\gamma}^{u}} - \frac{1}{u^2}(1 - e^{-j\Omega_{max_\gamma}^{u}}) \right.$$

$$\left. + \frac{\Omega_{max_\gamma}}{ju} e^{j\Omega_{max_\gamma}^{u}} + \frac{1}{u^2}(e^{j\Omega_{max_\gamma}^{u}} - 1)\right\}. \qquad (2.3.12b)$$

Collecting terms,

$$h_{\Omega_{max_\gamma}}(u) = \frac{1}{2\pi}\left\{ \frac{\Omega_{max_\gamma}}{ju} (e^{j\Omega_{max_\gamma}^{u}} - e^{-j\Omega_{max_\gamma}^{u}}) \right.$$

$$\left. + \frac{1}{u^2}[e^{j\Omega_{max_\gamma}^{u}} + e^{-j\Omega_{max_\gamma}^{u}} - 2]\right\}. \qquad (2.3.13)$$

Using

$$\sin^2(a) = \left(\frac{1}{j2}\right)^2 [e^{ja} - e^{-ja}]^2 = \frac{-1}{4}[e^{j2a} + e^{-j2a} - 2] \qquad (2.3.14)$$

gives

$$h_{\Omega_{max_\gamma}}(u) = \frac{1}{2\pi}\{ (2\Omega_{max_\gamma}{}^2)\frac{\sin(\Omega_{max_\gamma}^{u})}{(\Omega_{max_\gamma}^{u})} - \Omega_{max_\gamma}^{2} (\tfrac{4}{4})\frac{\sin^2\left(\frac{\Omega_{max_\gamma}^{u}}{2}\right)}{\left(\frac{\Omega_{max_\gamma}^{u}}{2}\right)^2}\} \qquad (2.3.15)$$

or finally,

$$(73)h_{\Omega_{max_\gamma}}(u) = \frac{\Omega_{max_\gamma}{}^2}{2\pi} \{2\frac{\sin(\Omega_{max}u)}{(\Omega_{max_\gamma}^{u})} - \frac{\sin^2\left(\frac{\Omega_{max_\gamma}^{u}}{2}\right)}{\left(\frac{\Omega_{max_\gamma}^{u}}{2}\right)^2}\}. \qquad (2.3.16)$$

If exact Nyquist sampling of $p_\phi(u)$ is chosen, then the spacing Δu is

$$\Delta u = 2\pi/(2\Omega_{max_\gamma}) = \pi/\Omega_{max_\gamma} \qquad (2.3.17)$$

so Eq. (2.3.16) becomes (because $\Omega_{max_\gamma} = \pi/\Delta u$)

$$h_{\Omega_{max_\gamma}}(u) = \frac{\pi}{2(\Delta u)^2}\left\{2\frac{\sin\left(\frac{\pi u}{\Delta u}\right)}{\left(\frac{\pi u}{\Delta u}\right)} - \frac{\sin^2\left(\frac{\pi u}{2\Delta u}\right)}{\left(\frac{\pi u}{2\Delta u}\right)^2}\right\}. \tag{2.3.18}$$

Now examine

$$Q_\phi(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{p}_\phi(k_u)\tilde{h}_{\Omega_{max_\gamma}}(k_u)e^{jk_u u}\,dk_u$$

$$= \int_{-\infty}^{\infty} p_\phi(u')h_{\Omega_{max_\gamma}}(u-u')du'. \tag{2.3.19}$$

The function $h_{\Omega_{max_\gamma}}(u)$ is by its definition bandlimited (see Eq. (2.3.9)), and it has been mentioned before (by invoking the finite support of $\tilde{h}_{\Omega_{max_\gamma}}(k_u)$) that $p_\phi(u)$ is also bandlimited because γ is. Therefore, these two functions can be expanded over sinc bases (see Section 3.1e). If the exact Nyquist sampling rate is used,

$$p_\phi(u) = \sum_{l=-\infty}^{\infty} p_\phi(l\Delta u)\cdot\frac{\sin[\Omega_{max_\gamma}(u-l\Delta u)]}{[\Omega_{max_\gamma}(u-l\Delta u)]} \tag{2.3.20}$$

$$h_{\Omega_{max_\gamma}}(u) = \sum_{n=-\infty}^{\infty} h(n\Delta u)\cdot\frac{\sin[\Omega_{max_\gamma}(u-n\Delta u)]}{[\Omega_{max_\gamma}(u-n\Delta u)]}. \tag{2.3.21}$$

(For oversampling other than exact Nyquist sampling, all Ω_{max_γ}'s in these equations would be replaced by $\pi/\Delta x$'s.) In Eq. (2.3.21) the samples of h are obtained from Eq. (2.3.18):

$$h(n\Delta u) = \frac{\pi}{2(\Delta u)^2}\left\{2\frac{\sin(\pi n)}{\pi n} - \frac{\sin^2\left(\frac{\pi n}{2}\right)}{\left(\frac{\pi n}{2}\right)^2}\right\} = \frac{\pi}{2(\Delta u)^2}\begin{cases} 2-1=1 & n=0 \\ 0 & n \text{ even} \\ -4/(\pi n)^2 & n \text{ odd} \end{cases}$$

$$= \frac{\pi}{2(\Delta u)^2}\begin{cases} 1 & n=0 \\ 0 & n \text{ even} \\ -(2/\pi n)^2 & n \text{ odd}. \end{cases} \tag{2.3.22}$$

(The subscript Ω_{max_γ} is now dropped for convenience, but the connection with the bandwidth of the object function still exists via the Nyquist constraint on Δu.) It now becomes necessary to invoke the orthogonality of the sinc functions.

2.3a Orthogonality of the sinc functions

The following is a demonstration of orthogonality of the sinc functions, and is based on the derivation given by McNamee et al., (1971). Beginning with the property

$$f(x-x') \Leftrightarrow e^{-jkx'}\tilde{f}(k) \tag{2.3a.1}$$

and, because

$$\frac{1}{\Delta x}\,\mathrm{sinc}\left(\frac{x}{\Delta x}\right) \Leftrightarrow \tilde{b}_{\frac{\pi}{\Delta x}}(k) \tag{2.3a.2}$$

where

$$\tilde{b}_{\frac{\pi}{\Delta x}}(k) = \begin{cases} 1 & |k| < \frac{\pi}{\Delta x} \\ 0 & \text{otherwise} \end{cases} \tag{2.3a.3}$$

then

$$\frac{1}{\Delta x}\mathrm{sinc}\left(\frac{x-x'}{\Delta x}\right) \Leftrightarrow e^{-jkx'}\tilde{b}_{\frac{\pi}{\Delta x}}(k), \tag{2.3a.4}$$

and orthogonality can be shown by using Parseval's relation.

The real-valued function version of Parseval's theorem is shown here in the form given by (Gabel and Roberts, 1980). From the convolution theorem for any real-valued functions g and h,

$$g(x)h(x) \Leftrightarrow \frac{1}{2\pi}\,\tilde{g}(k)*\tilde{h}(k). \tag{2.3a.5}$$

Then the left-hand side of Parseval's relation can be written as

$$\int_{-\infty}^{\infty} g(x)h(x)dx = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\{\int_{-\infty}^{\infty}[\int_{-\infty}^{\infty}\tilde{g}(k')\tilde{h}(k-k')dk']e^{jkx}dk\}dx. \tag{2.3a.6}$$

Noting that

$$\delta(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{jkx}dx \tag{2.3a.7}$$

and carrying out the integration over x in Eq. (2.3a.6),

$$\int_{-\infty}^{\infty} g(x)h(x)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\delta(k)\int_{-\infty}^{\infty}\tilde{g}(k')\tilde{h}(k-k')dk'dk \tag{2.3a.8}$$

so that

$$\int_{-\infty}^{\infty} g(x)h(x)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{g}(k)\tilde{h}(-k)dk. \tag{2.3a.9}$$

Parenthetically, if g(x) and h(x) are now taken to be complex functions then, using the fact that

$$f^*(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}^*(k)e^{-jkx}dk = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{f}^*(-k)e^{jkx}dk \tag{2.3a.10}$$

implies $f^*(x) \Leftrightarrow \tilde{f}^*(-k)$, then from Eq. (2.3a.9) the complex form of Parseval's theorem is obtained:

$$\int_{-\infty}^{\infty} g(x)h^*(x)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{g}(k)\tilde{h}^*(k)dk. \tag{2.3a.11}$$

Applying Eq. (2.3a.9) to the case at hand,

$$\int_{-\infty}^{\infty}\mathrm{sinc}(\tfrac{x-s}{\Delta x})\,\mathrm{sinc}(\tfrac{x-t}{\Delta x})dx = \frac{(\Delta x)^2}{2\pi}\int_{\frac{-\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{-jk(s-t)}dk = \Delta x\,\mathrm{sinc}(\tfrac{s-t}{\Delta x}) \tag{2.3a.12}$$

where the last equality followed from Eq. (2.3a.2). If now $s = m\Delta x$ and $t = n\Delta x$ where m and n are integers,

$$\frac{1}{\Delta x}\int_{-\infty}^{\infty}\mathrm{sinc}(\tfrac{x-m\Delta x}{\Delta x})\mathrm{sinc}(\tfrac{x-n\Delta x}{\Delta x})dx = \mathrm{sinc}(m-n) = \begin{cases} 0 & m\neq n \\ 1 & m=n \end{cases} \tag{2.3a.13}$$

because sinc(m) equals a kroneker delta function.

Thus, the convolution in Eq. (2.3.19) reduces to an infinite sum:

$$Q_\phi(m\Delta u) = \Delta u \sum_{l=-\infty}^{\infty} p_\phi(l\Delta u)h[(m-l)\Delta u]. \qquad (2.3a.14)$$

If $p_\phi(l\Delta u) = 0$ for l outside the range (0,N - 1) (due to the diminishing amplitude of the scattered field at wide angles from the center-line ray) then Q_ϕ becomes

$$Q_\phi(m\Delta u) \simeq \Delta u \sum_{l=0}^{N-1} p_\phi(l\Delta u)h[(m-l)\Delta u] \quad 0 \leq m \leq N-1 \qquad (2.3a.15)$$

where now h(·) is identified as the filter of the "filtered backprojections." The convolution in Eq. (2.3a.15) is aperiodic, so if it is carried out in the discrete Fourier transform domain via the FFT, h and p_ϕ must be zero-padded out to 2N - 1 elements. Then

$$Q_\phi(m\Delta u) = \Delta u \mathrm{FFT}^{-1}\{\underbrace{\mathrm{FFT}(p_\phi \text{ with zero padding}) \cdot \mathrm{FFT}(h \text{ with zero-padding})}_{A} \cdot (\text{smoothing window if desired})\}. \qquad (2.3a.16)$$

In Eq. (2.3.5), a continuum of viewing angles ϕ will not be available. Rather, there will be a number I of discrete viewing angles ϕ_i, so that, with π/I as the angular increment, Eq. (2.3.7) becomes

$$\gamma(x,y) \simeq \frac{\pi}{2\pi I} \sum_{i=1}^{I} Q_\phi$$

$$= \frac{1}{2I} \sum_{i=1}^{I} Q_{\phi\, i}(x\cos\phi_i + y\sin\phi_i). \qquad (2.3a.17)$$

Note that the values of u as arguments to Q_{ϕ_i} may not be exactly $m\Delta u$. So a linear interpolation of the Q_{ϕ_i} must be used to estimate Q_{ϕ_i} for such $u \neq m\Delta u$. So here again one has an interpolation problem--just what was trying to be avoided! But it is easily handled by simply zero-padding A in Eq. (2.3a.16) for superresolution as much as necessary.

The reasons for the words "convolution" or "filtered" in the name of this algorithm should now be clear. The reason for the word "backprojection" is as follows. To every point (x,y) in the image plane there corresponds a value of u for a given value of ϕ_i. The contribution that Q_{ϕ_i} makes to the reconstruction at location (x,y) is its value for the corresponding value of u. In Fig. (2.3a.1), for each point on the line, Q_{ϕ_i} will make the same contribution because u is constant. In this way Q_ϕ is "smeared" or "backprojected" over the image plane. The sum of all such smearings results in the reconstruction image.

2.4 The Filtered Backpropagation Algorithm

In conventional tomography, one begins with $\tilde{\gamma}(k_u,\phi)$ known on a discrete set of radial lines corresponding to discrete viewing angles ϕ_i. In diffraction tomography one begins with

$$\tilde{\gamma}[k_0(\hat{s}-\hat{s}_0)] = \tilde{\gamma}[k_u,-(k_0-k_v^*)] \qquad (2.4.1)$$

known on a discrete set of semicircles corresponding to discrete viewing angles ϕ_i. Devaney (1982), recognizing that reconstructions using convolution backprojection were superior to Fourier-domain interpolation algorithms, sought to apply the same concept to the diffraction tomography problem. In conventional tomography, the convolution backprojection theorem led from

$$\tilde{p}_\phi(k_u) = \tilde{\gamma}(k_u,\phi) \tag{2.4.2}$$

to

$$\gamma(x,y) = \frac{1}{2\pi}\int_0^\pi \{\frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{p}_\phi(k_u)|k_u| e^{jk_u u} dk_u\} d\phi. \tag{2.4.3}$$

The task in diffraction tomography was to go from

$$\tilde{D}_\phi(k_u,\omega) = \tilde{\gamma}[k_0(\hat{s}-\hat{s}_0)] \tag{2.4.4}$$

(the function $\tilde{D}_\phi$ being defined in Eqs. (2.2c.18) and 2.2c.19)) to an equation analogous to Eq. (2.3.5) (same as(2.4.3)). As was necessary in conventional convolution backprojection, it is necessary to assume that γ is bandlimited to $|\tilde{k}| < \Omega_{max_\gamma}$. From Fig. (2.1.1) it is evident that $\hat{s}_0 = (\cos\phi_0, \sin\phi_0)$ because $\hat{s}_0 = \hat{v}$ is the direction of the incident plane wave (see Eq. (2.2c.11)). From Eq. (2.2c.11), $\hat{s}$ is also a unit vector, pointing in a different direction, called χ by Devaney (1982). In Fig. (2.4.1) $\hat{s}$, $\hat{s}_0$, and χ are shown in relation to the k_x and k_y axes. In Eq. (2.2c.11) $(k_x,k_y) = k_0(\hat{s} - \hat{s}_0)$ was written. Now that relationship is made explicit with respect to χ and ϕ_0:

$$\hat{s}_0 \overset{(a)}{=} (\cos\phi_0, \sin\phi_0)$$
$$\hat{s} \overset{(b)}{=} (\cos\chi, \sin\chi)$$
$$k_x \overset{(c)}{=} k_0(\cos\chi - \cos\phi_0)$$
$$k_y \overset{(d)}{=} k_0(\sin\chi - \sin\phi_0). \qquad (2.4.5)$$

Using the same method as that used in conventional tomography, one begins by writing

$$\gamma(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{\gamma}(k_x,k_y)e^{j(k_x x + k_y y)}dk_x dk_y \qquad (2.4.6)$$

and trying to put the integral into polar form, while leaving the left-hand side evaluated in rectangular coordinates. Devaney (1982) d this in two steps. First the integration variables in Eq. (2.4.6) are changed to (ϕ_0,χ). This is accomplished via relations (2.4.5) and by forming the Jacobian

$$J = \det\begin{bmatrix} \frac{\partial k_x}{\partial\chi} = -k_0\sin\chi & \frac{\partial k_y}{\partial\chi} = k_0\cos\chi \\ \frac{\partial k_x}{\partial\phi_0} = k_0\sin\phi_0 & \frac{\partial k_y}{\partial\phi_0} = -k_0\cos\phi_0 \end{bmatrix} = k_0^2[\sin\chi\cos\phi_0 - \sin\phi_0\cos\chi]$$
$$= k_0^2\sin(\chi-\phi_0)$$
$$= k_0^2\sqrt{1 - \cos^2(\chi-\phi_0)}$$
$$= k_0^2\sqrt{1 - (\hat{s}\cdot\hat{s}_0)^2}. \qquad (2.4.7)$$

So

$$dk_x dk_y = Jd\chi d\phi_0 = k_0^2\sqrt{1 - (\hat{s}\cdot\hat{s}_0)^2}d\chi d\phi_0. \qquad (2.4.8)$$

Clearly, if γ is modeled as being bandlimited to $\Omega_{max_\gamma} = \sqrt{2}k_0$ (which is the range over which information about it is available, from the Fourier Diffraction Theorem), then the integration over the entire (k_x,k_y) plane reduces to integration over the $\sqrt{2}k_0$ circle. This is covered by allowing ϕ_0 and χ both to vary

from $-\pi$ to $+\pi$. Actually, this will cover the Fourier space twice, so the result must be divided by 2. (For example, both ($\phi_0 = \pi/4, \chi = \pi/6$) and ($\phi_0 = \pi/6, \chi = \pi/4$) are included.) Then (using Eq. (2.2c.19) for $\tilde{\gamma}$)

$$\gamma(x,y) = \frac{k_0^2}{2(2\pi)^2}\int_{-\pi}^{\pi} d\phi_0 \int_{-\pi}^{\pi} d\chi \sqrt{1 - (\hat{s}\cdot\hat{s}_0)^2}\, \tilde{\gamma}[k_0(\hat{s}-\hat{s}_0)]e^{jk_0(\hat{s}-\hat{s}_0)\cdot\vec{r}} \tag{2.4.9}$$

By the bandlimited model of γ,

$$\tilde{\gamma}[k_0(\hat{s}-\hat{s}_0)] = 0 \tag{2.4.10}$$

if

$$|k_0(\hat{s}-\hat{s}_0)| > \sqrt{2}k_0 \tag{2.4.11}$$

or if

$$(\hat{s}-\hat{s}_0)\cdot(\hat{s}-\hat{s}_0) > 2$$

$$1-2\hat{s}_0\cdot\hat{s}+1-2 > 0$$

$$\hat{s}_0\cdot\hat{s} < 0. \tag{2.4.12}$$

For convenience, now redefine χ to be the angle between $\hat{s}$ and the k_u axis rather than the k_x axis (see Fig. (2.4.2)). From Eq. (2.2c.11),

$$\hat{s}_0\cdot\hat{s} = \frac{k_v^*}{k_0} = \frac{\sqrt{k_0^2 - k_u^2}}{k_0} \geq 0 \tag{2.4.13}$$

if $0 \leq \chi \leq \pi$ (see Fig. (2.4.2)) and $\hat{s}_0\cdot\hat{s} < 0$ for $-\pi \leq \chi \leq 0$, for which case, as noted above, $\tilde{\gamma} \simeq 0$ has been assumed. Therefore, χ need range in Eq. (2.4.9) only from 0 to π:

$$\chi = 0 \Leftrightarrow k_u = k_0$$
$$\chi = \pi \Leftrightarrow k_u = -k_0$$
$$\int_0^{\pi} d\chi \Leftrightarrow \int_{k_0}^{-k_0} dk_u = -\int_{-k_0}^{k_0} dk_u. \tag{2.4.14}$$

Now

$$k_u = k_0\cos\chi \rightarrow -\sin\chi d\chi = \frac{dk_u}{k_0} . \tag{2.4.15}$$

But

$$k_v^* = k_0\sqrt{1 - (\frac{k_u}{k_0})^2} = k_0\sin\chi \rightarrow d\chi = -\frac{dk_u}{k_v^*} . \tag{2.4.16}$$

Also

$$\sqrt{1 - (\hat{s}\cdot\hat{s}_0)^2} = \sqrt{1 - (\frac{\sqrt{k_0^2 - k_u^2}}{k_0})^2} = \sqrt{1 - 1 + (\frac{k_u}{k_0})^2} = \frac{|k_u|}{k_0} . \tag{2.4.17}$$

By Eq. (2.2c.11),

$$k_0(\hat{s}-\hat{s}_0)\cdot\vec{r} = k_u u + (k_v^* - k_0)v \tag{2.4.18}$$

$$\gamma(x,y) = -(\frac{-k_0^2}{2(2\pi)^2})\int_{-\pi}^{\pi} d\phi_0 \int_{-k_0}^{k_0} \frac{dk_u}{k_v^*} \frac{|k_u|}{k_0} \tilde{\gamma}(k_u, k_v^*-k_0) e^{j[k_u u+(k_v^*-k_0)v]} \tag{2.4.19}$$

From Eq. (2.2c.19),

$$\gamma(x,y) = \frac{k_0}{2(2\pi)^2}\int_{-\pi}^{\pi} d\phi_0 \int_{-k_0}^{k_0} \tilde{D}_\phi(k_u,\omega) \frac{e^{j[k_u u+(k_v^*-k_0)v]}}{k_v^*} |k_u| dk_u . \tag{2.4.20}$$

Equation (2.4.20) is the equation analogous to Eq. (2.3.5) in straight-path tomography. Now substituting the expression for $\tilde{D}_\phi$ in Eq. (2.2c.25) into Eq. (2.4.20) and collecting additive terms in the exponential, Eq. (2.4.20) becomes

$$\gamma(x,y) = \frac{-jk_0}{(2\pi)^2}\int_{-\pi}^{\pi} d\phi_0 \int_{-k_0}^{k_0} dk_u |k_u| \tilde{\Gamma}_\phi(k_u,\omega) e^{j[k_u u+(k_v^*-k_0)(v-L)]}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \{\frac{-jk_0}{2\pi}\int_{-k_0}^{k_0} dk_u |k_u| \tilde{\Gamma}_\phi(k_u,\omega) e^{j[k_u u+(k_v^*-k_0)(v-L)]}\} d\phi_0 . \tag{2.4.21}$$

Define $\hat{Q}_{\phi_0}(u)$ as $\Gamma_\phi(u,\omega)$ filtered by $h_{k_0}(u)$ of Eq. (2.3.16), or, specified in the Fourier domain by $\tilde{h}_{k_0}(k_u)$ in Eq. (2.3.9). Specifically,

$$Q_\phi(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} |k_u| \begin{bmatrix} 1 & |k_u| \le k_0 \\ 0 & |k_u| > k_0 \end{bmatrix} \tilde{\Gamma}_\phi(k_u,\omega) e^{jk_u u} dk_u \qquad (2.4.22)$$

so that, within the range of k_u in Eq. (2.4.21), $|k_u|\tilde{\Gamma}_\phi(k_u,\omega)$ can be replaced by

$$\int_{-\infty}^{\infty} du' e^{-jk_u u'} Q_{\phi_0}(u'). \qquad (2.4.23)$$

Note that the exponential in Eq. (2.4.21), $e^{j\{k_u u+(k_v^*-k_0)(v-L)\}}$ is the essential difference from conventional tomographic reconstruction using filtered backprojection (see, for example, Eq. (2.3.5)). To turn this expression into a form with explicit (x,y) dependence on the right-hand side, Eq. (2.4.21) is rewritten as

$$\gamma(x,y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \{-jk_0 \int_{-\infty}^{\infty} du' Q_\phi(u') \cdot \frac{1}{2\pi}\int_{-k_0}^{k_0} dk_u e^{j[k_u(u-u')+(k_v^*-k_0)(v-L)]}\} d\phi. \qquad (2.4.24a)$$

The term from the second $1/2\pi$ to the right curly bracket is identified as $G(u-u',v-L)$, so

$$\gamma(x,y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \{-jk_0 \int_{-\infty}^{\infty} du' Q_\phi(u') G(u-u',v-L)\} d\phi \qquad (2.4.24b)$$

where

$$G(u,v) = \frac{1}{2\pi}\int_{-k_0}^{k_0} dk_u e^{j[k_u u+(k_v^*-k_0)v]} \qquad (2.4.25)$$

where the term in curly brackets will be defined as $\Pi_{\phi_0}(u,v)$. But

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x\sin\phi_0 - y\cos\phi_0 \\ x\cos\phi_0 + y\sin\phi_0 \end{pmatrix}. \tag{2.4.26}$$

So

$$\gamma(x,y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Pi_{\phi_0}(x\sin\phi_0 - y\cos\phi_0,\ x\cos\phi_0 + y\sin\phi_0)\,d\phi_0. \tag{2.4.27}$$

2.5 Computational Complexity

Compared with Eq. (2.3.7), it is clear that Eq. (2.4.27) has an added dimension of complexity due to the two-dimensional propagator exponential. That is, $\Pi_{\phi_0}(u,v)$ must be computed for both u and v, while in Eq. (2.3.5) $Q_\phi(u)$ needs to be calculated only for different u values. The reason for this is that the $\hat{Q}_\phi(u)$ functions backpropagate to each point in the object region distinctly, while $Q_\phi(u)$ contribute only to points along a single line passing through the object region. Thus, $\Pi_{\phi_0}(u,v)$ is filtered with $|k_u|$ just as is $p_\phi(u)$ to obtain $\hat{Q}_\phi(u)$ and $Q_\phi(u)$, respectively. The difference comes with the additional filtering by G(u,v) to obtain Π_{ϕ_0}. The $\Pi_{\phi_0}(f_1(x,y),f_2(x,y))$ where

$$\begin{aligned} f_1(x,y) &\overset{(a)}{=} x\sin\phi_0 - y\cos\phi_0 = x\cos\phi + y\sin\phi = u \\ f_2(x,y) &\overset{(b)}{=} x\sin\phi + y\cos\phi = v \end{aligned} \tag{2.5.1}$$

is summed analogously to the $Q_\phi(f_1(x,y))$ in Eq. (2.3.7) to obtain the final image of $\gamma(x,y)$. To evaluate Eq. (2.4.27) and estimate its computational complexity, refer to (Pan and Kak, 1983). There Eq. (2.4.21) is written as

$$\gamma(x,y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \{\frac{-jk_0}{2\pi}\int_{-\infty}^{\infty} \tilde{\Gamma}_\phi(k_u)\tilde{h}(k_u)\tilde{G}_v(k_u)e^{jk_u u}dk_u\}d\phi_0 \tag{2.5.2}$$

where

$$\tilde{h}(k_u) = \begin{cases} |k_u| & |k_u| \le k_0 \\ 0 & \text{otherwise} \end{cases} \qquad (2.5.3)$$

and

$$\tilde{G}_v(k_u) = e^{j[(k_v^*-k_0)(v-L)]} \qquad (2.5.4)$$

where without $\tilde{G}_v(k_u)$, Eq. (2.5.2) is essentially identical to Eq. (2.3.5). It is clear that if there are N_v depths chosen in the object region (see Fig. (2.5.1)) N_v FFTs will need to be done for a single projection. If there are N_ϕ projections, then N_vN_ϕ FFT s are required. An FFT takes the order of $N \cdot \log(N)$ operations, where N is the number of sample points on a projection. For an nxn reconstruction, typically $N_\phi \sim N_v \sim N \sim n$, so the total number of operations per reconstruction is of the order $n^3\log(n)$. Obviously, because the only computationally significant difference between filtered backpropagation and filtered backprojection is the depth-dependent $\tilde{G}_v(k_u)$, for filtered backprojection the computational complexity is of the order $n^2\log(n)$ because there is no $N_v \sim n$ factor (one $\tilde{Q}_\phi(k_u)$ is good for all depths v). For the case of Fourier domain interpolation and inversion, for an nxn reconstruction the interpolation (order n^2) is less than that of the two-dimensional FFT inversion, $n^2\log(n)$, so the computational complexity is the same as that of the convolution backprojection: $n^2\log(n)$.

CHAPTER 3

THE SINC BASIS MOMENT METHOD

3.0 Introduction

Having discussed several conventional tomography algorithms and having compared their ranges of validity and their orders of computation, attention is now turned to the focus of this thesis: the sinc basis moment method. In a recent series of papers [Johnson and Tracy (1983), Tracy and Johnson (1983), and Johnson et al. (1984)], discrete/numerical solutions of the ultrasonic scattering and inverse scattering problem, described by the exact inhomogeneous Helmholtz wave equation, were presented which under certain conditions yield quantitative spatial distributions of speed of sound and absorption. Perturbation approximations were not introduced. Although for a given density of sampling there are still limitations on the degree of contrast of an object for which the solution algorithm converges, the reconstructions are much better than those of the first Born approximation.

In the study considered in this chapter, effects of varying the sampling density, transducer-to-object region center distance, object contrast and radius, object grid size, and iteration duration are investigated. First, Johnson and Tracy's (1983) formulation and solution algorithm are restated. Miscellaneous programming details and issues are discussed next. In particular, the topics of scattered field generation and sampling are examined. Also, numerous details of theory necessary to comprehend in order to understand and evaluate this simulation study

are given, including in particular a derivation of the wave equation relevant to this thesis, a discussion of the Algebraic Reconstruction Technique, and an analysis of the coefficients in the matrix equations and their computation. Finally, a summary of computational results presented in Section 3.3 indicates the utility of the sinc basis method for solving small inverse scattering problems, and the effects of varying several parameters. The specific case of an infinite circular cylindrical object is considered.

3.1 Problem Formulation

The inhomogeneous wave equation approximately describing the propagation of ultrasound in tissue is, assuming negligible spatial variation in density,

$$(\nabla^2 + k_0^2) f(\vec{r}) = -\gamma(\vec{r}) f(\vec{r}) \qquad (3.1.1)$$

where $k_o = \omega/c_o$ is the wavenumber in the reference medium, where ω is the radial frequency of the insonifying wave and c_o is the speed of sound in the reference medium, $f(\vec{r})$ is the total ultrasonic field, and

$$\gamma(\vec{r}) = \omega^2(1/c(\vec{r})^2 - 1/c_0^2) - j\omega\alpha(\vec{r})/c(\vec{r}) \qquad (3.1.2)$$

is the object function to be reconstructed from scattered field data, $\alpha(\vec{r})$ is the (spatially varying) power (twice the pressure) absorption coefficient, and $c(\vec{r})$ is the speed of sound distribution function in the object region.

3.1a Derivation of the linear wave equation from basic acoustics equations

This section gives a derivation, from the basic equations of acoustics, of the wave equation to be used throughout this study, Eq. (3.1.1). It is a first-order wave equation, ignoring several second-order terms in addition to density variations. Yet even a more complicated wave equation accounting for density variations appearing in (Johnson and Tracy, 1983) still ignores second-order terms. The basic method is based on that by Chernov (1960), though some modifications have been made here. In addition, some comments about the first Born approximation will be convenient to make which complement those made in Chapter 2. First, the derivation will be given in the time domain, and then, again from the basic acoustic equations, in the temporal frequency domain. One begins by writing perturbation forms for pressure, density, and velocity as follows. The representation for pressure is

$$p = p_0 + p_1 \tag{3.1a.1}$$

where p is the total pressure, p_0 is the ambient, spatially and temporally constant background pressure, and p_1 is the pressure fluctuation. The symbol p is used for pressure in this subsection only, for the sake of convention and familiarity of appearance of the equations. (As noted in Section 2.2 and Chapter 1, f in Eq. (3.1.1) can represent one of several different scalar field quantities, depending upon the field quantity sought or the modality of tissue interrogation.) Similarly, for density fluctuations,

$$\rho = \rho_0 + \rho_1 \tag{3.1a.2}$$

where ρ is the total density, ρ_0 is the temporally constant ambient density, and ρ_1 is the density fluctuation. Finally, the analogous description of particle velocity is

$$\vec{v} = \vec{v}_0 + \vec{v}_1 \qquad (3.1a.3)$$

where the background partical velocity $\vec{v}_0 = 0$, so the total particle velocity $\vec{v}$ equals the velocity fluctuation $\vec{v}_1$. The equation of conservation of mass is

$$\frac{\partial \rho}{\partial t} \overset{(a)}{=} -\nabla\cdot(\rho\vec{v}) \quad \text{or} \quad \frac{\partial(\rho_0+\rho_1)}{\partial t} \overset{(b)}{=} -\nabla\cdot[(\rho_0+\rho_1)\vec{v}]. \qquad (3.1a.4)$$

Noting that $\partial\rho_0/\partial t = 0$ and ignoring the second-order term $\nabla\cdot(\rho_1\vec{v})$ gives

$$\frac{\partial \rho_1}{\partial t} = -\nabla\cdot\rho_0\vec{v}. \qquad (3.1a.5)$$

If in addition it is assumed that $\nabla\rho_0 = 0$ (the background density is spatially invariant), then

$$\frac{\partial \rho_1}{\partial t} = -\rho_0\nabla\cdot\vec{v} \qquad (3.1a.6)$$

the linearized equation of conservation of mass as usually stated. Differentiating Eq. (3.1a.5) with respect to time yields

$$\frac{\partial^2 \rho_1}{\partial t^2} = -\nabla\cdot(\rho_0\frac{\partial\vec{v}}{\partial t}) \qquad (3.1a.7)$$

where $\partial\rho_0/\partial t = 0$ was used. It should be noted that it is crucial to ignore the second-order term $\nabla\cdot(\rho_1\vec{v})$ in Eq. (3.1a.4b) because otherwise the time derivative would here yield another term, $\nabla\cdot[(\partial\rho_1/\partial t)\vec{v}]$, which would prohibitively complicate the

resulting wave equation.

The equation of motion can be written

$$\rho\frac{d\vec{v}}{dt} \overset{(a)}{=} -\nabla p \quad \text{or} \quad (\rho_0+\rho_1)\frac{d\vec{v}}{dt} \overset{(b)}{=} -\nabla(p_0+p_1). \tag{3.1a.8}$$

Noting that $\nabla p_0 = 0$ and ignoring the second-order terms $\rho_1(d\vec{v}/dt)$ and $(\vec{v}\cdot\nabla)\vec{v}$ gives

$$\rho_0\frac{\partial\vec{v}}{\partial t} = -\nabla p_1. \tag{3.1a.9}$$

In view of the linearization of the equation of conservation of mass discussed above, the second-order term $\rho_1(d\vec{v}/dt)$ must be ignored in Eq. (3.1a.8b) because in Eq. (3.1a.7) only ρ_0 appears, so that keeping $\rho_1(d\vec{v}/dt)$ would prohibitively complicate the wave equation resulting from combining Eqs. (3.1a.8b) and (3.1a.7). (Note that, if dealing only with the conservation of mass equation, one can, without added complexity, use the only partially linearized form

$$\rho\frac{\partial\vec{v}}{\partial t} = -\nabla p \tag{3.1a.10}$$

where only $(\vec{v}\cdot\nabla)\vec{v}$ has been left out.) Substituting $\rho_0(\partial\vec{v}/\partial t)$ in Eq. (3.1a.9) into Eq. (3.1a.7) gives

$$\frac{\partial^2\rho_1}{\partial t^2} = \nabla^2 p_1. \tag{3.1a.11}$$

For the equation of state one begins with the already linearized form

$$\frac{dp}{dt} \overset{(a)}{=} c^2 \frac{d\rho}{dt}$$

$$\frac{dp_1}{dt} \overset{(b)}{=} c^2 \frac{d\rho}{dt} \qquad (3.1a.12)$$

where the second equation holds because $\partial p_0/\partial t = \nabla p_0 = 0$ so that $dp_0/dt = 0$. Expanding d/dt on both sides in terms of the material derivatives (in order to obtain expressions with partial derivatives with respect to time) results in

$$\frac{1}{c^2}\left[\frac{\partial p_1}{\partial t} + \vec{v}\cdot\nabla p_1\right] = \frac{\partial(\rho_0+\rho_1)}{\partial t} + \vec{v}\cdot\nabla(\rho_0+\rho_1). \qquad (3.1a.13)$$

Again noting that $\partial\rho_0/\partial t = 0$ and ignoring the second-order terms $\vec{v}\cdot\nabla p_1$ and $\vec{v}\cdot\nabla\rho_1$ because not doing so would fundamentally complicate the resulting wave equation in the temporal domain, one obtains

$$\frac{1}{c^2}\frac{\partial p_1}{\partial t} = \frac{\partial \rho_1}{\partial t} + \vec{v}\cdot\nabla\rho_0. \qquad (3.1a.14)$$

The term $\vec{v}\cdot\nabla\rho_0$ is the source of an additional term in the wave equation, which accounts for spatial variations in the background density of the medium. Differentiating Eq. (3.1a.14) with respect to time results in the wave equation:

$$\text{(a)}\quad \frac{1}{c^2}\frac{\partial^2 p_1}{\partial t^2} = \frac{\partial^2 \rho_1}{\partial t^2} + \frac{\partial \vec{v}}{\partial t}\cdot\nabla\rho_0$$

$$\text{(b)}\quad \frac{1}{c^2}\frac{\partial^2 p_1}{\partial t^2} - \nabla^2 p_1 + \nabla(\ln \rho_0)\cdot\nabla p_1 = 0 \qquad (3.1a.15)$$

where Eq. (3.1a.11) has been used for $\partial^2\rho_1/\partial t^2$, Eq. (3.1a.9) for $\partial\vec{v}/\partial t$, and ln denotes natural logarithm. Dropping indices and

letting p represent acoustic pressure,

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} - \nabla^2 p + \nabla(\ln \rho)\cdot\nabla p = 0 \qquad (3.1a.16a)$$

where now ρ is the background density. If at this stage $e^{j\omega t}$ time dependence is assumed for p,

$$(\nabla^2 + k^2)p = \nabla(\ln\rho)\cdot\nabla p \qquad (3.1a.16b)$$

which is in agreement with (Johnson and Tracy, 1983).

Next the wave equation is derived from basic acoustic equations, immediately assuming harmonic time dependence (for all time-varying functions) of frequency ω. The steps are the same as above, but this time the total density will be explicitly kept until the end. In the equation of conservation of mass, this time only $\partial\rho_0/\partial t = 0$ is needed to obtain

$$j\omega\rho_1 = [-\nabla\cdot(\rho_0+\rho_1)\vec{v}]. \qquad (3.1a.17)$$

Recalling that the product of two functions of time dependence $e^{j\omega t}$ is another function with time dependence $e^{j2\omega t}$, differentiation of Eq. (3.1a.17) with respect to time gives

$$-\omega^2\rho_1 = -\nabla\cdot(\rho_0 j\omega\vec{v} + \rho_1 j2\omega\vec{v}). \qquad (3.1a.18)$$

The equation of motion in the temporal frequency domain is

$$\rho[j\omega\vec{v} + (\vec{v}\cdot\nabla)\vec{v}] = -\nabla p. \qquad (3.1a.19)$$

Again using $\nabla p_0 = 0$ and ignoring $(\vec{v}\cdot\nabla)\vec{v}$ yields

$$j\omega\rho\vec{v} = -\nabla p_1 \qquad (3.1a.20a)$$

or

$$j\omega\rho_0\vec{v} = -\nabla p_1 - j\omega\rho_1\vec{v}. \qquad (3.1a.20b)$$

Substituting Eq. (3.1a.20b) into Eq. (3.1a.18) results in

$$-\omega^2\rho_1 = \nabla^2 p_1 - j\omega\nabla\cdot(\rho_1\vec{v}). \qquad (3.1a.21)$$

From Eq. (3.1a.13), the equation of state can be written

$$\frac{1}{c^2}(j\omega p_1 + \vec{v}\cdot\nabla p_1) = j\omega\rho_1 + \vec{v}\cdot\nabla(\rho_0+\rho_1). \qquad (3.1a.22)$$

Differentiating Eq. (3.1a.22) with respect to time, and substituting Eq. (3.1a.21) for $-\omega^2\rho_1$,

$$-(\tfrac{\omega}{c})^2 p_1 + \frac{j2\omega\vec{v}}{c^2}\cdot\nabla p_1 = -\omega^2\rho_1 + j\omega(\vec{v}\cdot\nabla\rho_0 + 2\vec{v}\cdot\nabla\rho_1) =$$

$$\nabla^2 p_1 - j\omega\rho_1(\nabla\cdot\vec{v}) + j\omega\vec{v}\cdot\nabla\rho_0 + j\omega\vec{v}\cdot\nabla\rho_1 \qquad (3.1a.23a)$$

or

$$(\nabla^2 + k^2)p_1 = 2j\frac{\omega\vec{v}}{c^2}\cdot\nabla p_1 + j\omega\rho_1(\nabla\cdot\vec{v}) - j\omega\vec{v}\cdot[\nabla(\rho_0+\rho_1)] =$$

$$j\omega[\vec{v}\cdot(\frac{2\nabla p_1}{c^2} - \nabla\rho) + \rho_1\nabla\cdot\vec{v}]. \qquad (3.1a.23b)$$

From Eq. (3.1a.20a),

$$(\nabla^2 + k^2)p_1 = -\frac{\nabla p_1}{\rho}\cdot[\frac{2\nabla p_1}{c^2} - \nabla\rho] - \rho_1\nabla\cdot(\frac{\nabla p_1}{\rho}). \qquad (3.1a.24)$$

If $\rho_1 << \rho_0$, then Eq. (3.1a.24) becomes

$$(\nabla^2 + k^2)p_1 = -\frac{\nabla p_1}{\rho_0}\cdot(\frac{2\nabla p_1}{c^2} - \nabla\rho) - \rho_1\nabla\cdot(\frac{\nabla p_1}{\rho_0}). \qquad (3.1a.25)$$

The only term in this equation that is first order is $(\nabla\rho_0/\rho_0)\nabla p_1$. When the other three (second-order) terms are ignored, Eq. (3.1a.16b) results.

Parenthetically, if all terms within the linearization of the equation of state are kept, one more term than was kept in Eq. (3.1a.24)--the term $(\vec{v}\cdot\nabla)\vec{v}$ in Eq. (3.1a.19)--will appear on the right-hand side of Eq. (3.1a.24):

$$-\nabla\cdot[\rho(\vec{v}\cdot\nabla)\vec{v}] = \nabla\cdot\{\rho[\nabla\cdot(\frac{\nabla p_1}{\rho})]\frac{\nabla p_1}{\omega^2\rho}\} \qquad (3.1a.26)$$

(using Eq. (3.1a.20a)), which is again at most a second-order term.

Note that Eq. (3.1a.16b), appearing in Johnson and Tracy (1983), is only a first-order wave equation; if in addition, the term on the right-hand side $\nabla(\ln\rho)\nabla p$ is neglected, the result is the wave equation used for reconstruction by practically all diffraction tomography studies (including this study):

$$(\nabla^2 + k^2)p = 0. \qquad (3.1a.27)$$

Even if $\nabla(\ln\rho)\nabla p$ were kept, in order to go from Eq. (3.1a.24) to Eq. (3.1a.16b), all temporal fluctuations in density about the background density (though not spatial variations in background density) had to be ignored. In Chernov's (1960) description of the Born approximation applied to the wave equation, he reintroduces a density spatial fluctuation $\Delta\rho$ about the mean value ρ_0. In the present tomography study, all variations in density, temporal and spatial are ignored, as discussion begins with Eq. (3.1a.27). So in the present parallel discussion of the Born approximation as applied to the wave equation only speed of sound fluctuations (possibly complex) will be considered. Furthermore, Chernov (1960) works in the time domain, and here the temporal frequency domain is used. Finally, Chernov (1960) considers only the case of an incident plane wave; here any unperturbed solution of the homogeneous wave equation

$$(\nabla^2 + k_0^2)p_0 = 0 \qquad (3.1a.28)$$

where $k_0 = \omega/c_0$ will suffice. Thus, following Chernov (1960), let $c = c_0 + \Delta c$ where c and Δc are both functions of position. Then

$$c^2 = c_0^2 + \Delta c^2 + 2c_0\Delta c. \qquad (3.1a.29)$$

By long division,

$$\frac{1}{c^2} \simeq \frac{1}{c_0^2} - \frac{2\Delta c}{c_0^3}. \qquad (3.1a.30)$$

Substituting into Eq. (3.1a.27),

$$(\nabla^2 + k_0^2)p = \frac{2\omega^2\Delta c}{c_0^3}p. \qquad (3.1a.31)$$

Recall that in the equations of conservation of mass and motion, p stood for the sum of background pressure p_0 and acoustic pressure p_1, and that in Eq. (3.1a.16b) (and Eq. (3.1a.27)), p stood for acoustic pressure. If now the acoustic pressure is decomposed into an unperturbed acoustic pressure p_0 propagating in an infinite homogeneous medium and a perturbed acoustic pressure p_1 (perturbed via spatial inhomogeneities in complex speed of sound), by definition p_0 will satisfy Eq. (3.1a.28) so that Eq. (3.1a.31) simplifies to

$$(\nabla^2 + k_0^2)p_1 = \frac{2\omega^2\Delta c}{c_0^3}(p_0 + p_1). \qquad (3.1a.32)$$

Here the term $(\omega^2 \cdot 2\Delta c/c_0^3)p_1$ is second order. The first Born approximation ignores this term, with the result

$$(\nabla^2 + k_0^2)p_1 = \frac{2\omega^2\Delta c}{c_0^3}p_0. \qquad (3.1a.33)$$

Actually, one could use the form

$$(\nabla^2 + k_0^2)p_1 = -\omega^2(\frac{1}{c^2} - \frac{1}{c_o^2})p \qquad (3.1a.34)$$

and consider that the term $-\omega^2(1/c^2 - 1/c_0^2)p_1$ is second order and ignore it; that would be the first Born approximation as typically used in diffraction tomography. Note that the identification of f_*^{sc} with f^{sc} in Chapter 2 is easily seen to be equivalent to the Born approximation as derived in this section where here p_1 and f^{sc} are analogous.

Parenthetically, as shown in Section 2.2, the form of Eq. (3.1.1) is also valid for the case of nonuniform density, providing the meanings of f and γ are appropriately modified. Johnson and Tracy (1983) have given an algorithm for determining the density distribution from the present object distribution, γ. The method is quite lengthy and has not yet been demonstrated numerically, but in theory, then, distributions of sound speed, absorption, and density could be obtained from the sinc basis moment method.

3.1b Derivation of explicit form of γ

Here two derivations of γ will be given. In both, $e^{+j\omega t}$ time dependence is assumed. The first begins by defining a complex wavenumber, which is substituted into the expression for a plane wave. Thus, let $k = k' - jk''$. An isotropic medium is assumed, so k is the same for all directions. Then

$$p = Ae^{-jk\hat{s}\cdot\vec{r}} \quad (a)$$
$$= Ae^{-j(k'-jk'')\hat{s}\cdot\vec{r}} \quad (b)$$
$$= Ae^{-(k''+jk')\hat{s}\cdot\vec{r}} \quad (c) \qquad (3.1b.1)$$

where p is the pressure complex amplitude at frequency ω and position $\vec{r}$, $\hat{s}$ is the direction of propagation of the plane wave, and A is a real or complex constant. To relate the real and imaginary parts of k to speed of sound and absorption, respectively, compute $|p|^2$ as

$$pp^* = |A|^2 e^{-2k''\hat{s}\cdot\vec{r}} \qquad (3.1b.2)$$

and then identify

$$k' = \omega/c \quad (a)$$
$$2k'' = \alpha \quad (b) \qquad (3.1b.3)$$

where c is the sound speed and α is the power absorption coefficient of the medium. Now examine ∇^2 of the plane wave propagating in a (locally) homogeneous space. The pressure p satisfies $(\nabla^2 + k^2)p = 0$ with the following identifications:

$$\nabla^2 p = (-jk)^2 p = -k^2 p$$
$$= -(k'^2 - k''^2 - 2jk'k'')p$$
$$= -[(\omega/c)^2 - k''^2 - j2(\omega/c)k'']p. \qquad (3.1b.4)$$

But $2k'' = \alpha$, so

$$\nabla^2 p = -[(\omega/c)^2 - (\alpha/2)^2 - j\omega(\alpha/c)]p. \qquad (3.1b.5)$$

For low absorption cases, which are the norm for tissue, the $(\alpha/2)^2$ term is negligible compared with $(\omega/c)^2$ and can be ignored. Now referencing this particular medium with respect to a

designated lossless homogeneous medium with wavenumber k_0 and speed of sound c_0, the wave equation can be written

$$(\nabla^2 + k_0^2)p = -(k^2 - k_0^2)p \simeq -[(\omega/c)^2 - (\omega/c_0)^2 - j\omega(\alpha/c)]p$$
$$= -\gamma p \qquad (3.1b.6)$$

where

$$\gamma \triangleq \omega^2\left(\frac{1}{c^2} - \frac{1}{c_0^2}\right) - j\omega\alpha/c. \qquad (3.1b.7)$$

For the second derivation, c and therefore k are complex quantities: $c = c' + jc''$ where k,c, c', c'' (and therefore α below) are all functions of $\vec{r}$. Then

$$k = \frac{\omega}{c'+jc''} = \frac{\omega(c'-jc'')}{c'^2+c''^2}. \qquad (3.1b.8)$$

The value of c''^2 is now ignored, compared with c'^2, so that

$$k = \frac{\omega}{c'} - j\frac{\omega c''}{c'^2} = k'-jk'' \qquad (3.1b.9)$$

and one can identify

$$k' = \omega/c' \qquad (3.1b.10a)$$

and

$$2k'' = 2\left(\frac{\omega c''}{c'^2}\right) = \alpha. \qquad (3.1b.10b)$$

Rewriting k^2, in light of the first derivation,

$$k^2 = \frac{\omega^2}{c'^2-c''^2+j2c'c''} = \frac{\omega^2(c'^2-c''^2-j2c'c'')}{(c'^2-c''^2)^2 + (2c'c'')^2}$$
$$\simeq (\omega/c')^2 - j\frac{\omega^2 2c'c''}{c'^4}$$
$$= (\omega/c')^2 - j\left(\frac{2\omega c''}{c'^2}\right)\frac{\omega}{c'}$$
$$= (\omega/c')^2 - \frac{j\omega\alpha}{c'}. \qquad (3.1b.11)$$

Now similarly to the previous derivation, reference the wave equation to the lossless coupling medium with speed of sound c_0 and let $k_0 = \omega/c_0$. Then

$$(\nabla^2+k^2)p = 0 \rightarrow (\nabla^2 + k_0^2)p = -(k^2 - k_0^2)p$$
$$= [(\omega/c')^2 - (\omega/c_0)^2 - j\frac{\omega\alpha}{c'}]p$$
$$= -\gamma p \qquad (3.1b.12)$$

where γ is the same as that defined above. Note that the usual definition of speed of sound at location $\vec{r}$ is just $c'(\vec{r})$. Solving for $c'(\vec{r})$ and $\alpha(\vec{r})$ in terms of γ gives

$$c'(\vec{r}) = \frac{1}{\sqrt{\frac{\mathrm{Re}[\gamma(\vec{r})]}{\omega^2} + \frac{1}{c_0^2}}} \qquad (3.1b.13)$$

and then

$$\alpha(\vec{r}) = -\frac{c'(\vec{r})}{\omega}\,\mathrm{Im}[\gamma(\vec{r})]. \qquad (3.1b.14)$$

3.1c Integral representation of the scattering problem

The differential equation Eq. (3.1.1) is solved via the Green function method over a volume both containing the tissue and having a surface including the transmitting transducer surface. The volume inhomogeneity function, which is the product of the total field times the object function, i.e., γf, is unknown and under the integral sign, making inversion difficult. The form of the integral equation for the scattered field is

$$f^{sc}(\vec{r}) = f(\vec{r}) - f^{inc}(\vec{r}) = \int_{-\infty}^{\infty} \gamma(\vec{r}')f(\vec{r}')G(|\vec{r}-\vec{r}'|)d\vec{r}' \qquad (3.1c.1)$$

where $f^{sc}(\vec{r})$ is the scattered field, $f^{inc}(\vec{r})$ is the incident

field, and $G(\cdot)$ is the free-space Green function.

The two-dimensional inverse scattering problem (Eq. (3.1c.1)) takes the form

$$f^{sc}(x,y) = f(x,y) - f^{inc}(x,y)$$
$$= (-j/4)\int_{-\infty}^{\infty}\int \gamma(x',y')f(x',y')H_0^{(2)}[k_0\sqrt{(x-x')^2+(y-y')^2}]dx'dy' \tag{3.1c.2}$$

where $(-j/4)H_0^{(2)}(k_0r)$ is the free-space Green function in two dimensions (assuming time dependence $e^{+j\omega t}$).

3.1d Derivation of the two-dimensional free-space Green function

The following is a derivation of the two-dimensional free-space Green function, based on that given by Tyras (1969). The free-space Green function satisfies

$$(\nabla_{\vec{r}}^2 + k_0^2)G(\vec{r},\vec{r}') = -\delta(\vec{r}-\vec{r}'). \tag{3.1d.1}$$

Here in two dimensions, $\vec{r} = (x,y)$, and the geometry is as shown in Fig. (3.1d.1). Take the two-dimensional Fourier transform of Eq. (3.1d.1), where $\vec{k} = (k_x,k_y)$ and $|\vec{k}| = k$:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\nabla_{\vec{r}}^2+k_0^2)G(\vec{r},\vec{r}')e^{-j\vec{k}\cdot\vec{r}}d\vec{r} = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(\vec{r}-\vec{r}')e^{-j\vec{k}\cdot\vec{r}}d\vec{r} \tag{3.1d.2a}$$

or

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial^2 G(\vec{r},\vec{r}')}{\partial x^2} + \frac{\partial^2 G(\vec{r},\vec{r}')}{\partial y^2} e^{-j\vec{k}\cdot\vec{r}}dxdy$$
$$+ k_0^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(\vec{r},\vec{r}')e^{-j\vec{k}\cdot\vec{r}}dxdy = -e^{-j\vec{k}\cdot\vec{r}'} \tag{3.1d.2b}$$

The third term on the left-hand side of Eq. (3.1d.2b) equals $k_0^2G(\vec{k},\vec{r}')$. Consider the first term:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial^2 G(\vec{r},\vec{r}')}{\partial x^2} e^{-j\vec{k}\cdot\vec{r}}dxdy = \int_{-\infty}^{\infty} \{\frac{\partial G(\vec{r},\vec{r}')}{\partial x} e^{-j\vec{k}\cdot\vec{r}} \Big]_{-\infty}^{\infty}$$

$$+ jk_x \int_{-\infty}^{\infty} \frac{\partial G(\vec{r},\vec{r}')}{\partial x} e^{-j\vec{k}\cdot\vec{r}}dx\}dy$$

$$= \int_{-\infty}^{\infty} \{[\frac{\partial G(\vec{r},\vec{r}')}{\partial x} + jk_x G(\vec{r},\vec{r}')]e^{-j\vec{k}\cdot\vec{r}} \Big]_{-\infty}^{\infty} - k_x^2 \int_{-\infty}^{\infty} G(\vec{r},\vec{r}')e^{-j\vec{k}\cdot\vec{r}}dx\}dy. \quad (3.1d.3)$$

The integrated term evaluated at $\pm\infty$ is zero because G must satisfy the radiation condition (Sommerfeld, 1949). The last term on the right-hand side is $k_x^2 G(\vec{k},\vec{r}')$. Similarly, the second term in Eq. (3.1d.2b) will yield $-k_y^2 G(\vec{k},\vec{r}')$ so that, recalling that $k^2 = k_x^2 + k_y^2$, Eq. (3.1d.2b) becomes

$$G(\vec{k},\vec{r}') = \frac{e^{-j\vec{k}\cdot\vec{r}}}{k^2-k_0^2}. \quad (3.1d.4)$$

Then

$$G(\vec{r},\vec{r}') = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{e^{j\vec{k}\cdot(\vec{r}-\vec{r}')}}{k^2-k_0^2} d\vec{k}. \quad (3.1d.5)$$

Now express this result in terms of a well-known function--the zeroth-order Hankel function of the second kind. To do this, convert the above integral into its polar coordinates form.

$$\left.\begin{aligned} k_x &= k\cos\beta \\ k_y &= k\sin\beta \end{aligned}\right\} \rightarrow \quad dk_x dk_y = kdkd\beta \begin{cases} 0\le k<\infty \\ -\pi\le\beta\le\pi \end{cases}$$

$$x-x' = |\vec{r}-\vec{r}'|\cos\varphi \qquad \vec{k}\cdot(\vec{r}-\vec{r}') = |\vec{k}|\ |\vec{r}-\vec{r}'|\cos(\beta-\varphi)$$

$$y-y' = |\vec{r}-\vec{r}'|\sin\varphi \quad (3.1d.6)$$

so that

$$G(\vec{r},\vec{r}') = \frac{1}{(2\pi)^2}\int_0^{\infty} \frac{kdk}{k^2-k_0^2}\int_{-\pi}^{\pi} e^{j(k|\vec{r}-\vec{r}'|)\cos(\beta-\varphi)}. \quad (3.1d.7)$$

But for any φ,

$$J_0(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{jz\cos(\beta-\varphi)}d\beta \qquad (3.1d.8)$$

and

$$J_v(z) = \frac{1}{2}\,[H_v^{(1)}(z) + H_v^{(2)}(z)]. \qquad (3.1d.9)$$

Also,

$$H_v^{(1)}(z) = -e^{-jv\pi}H_v^{(2)}(-z) \qquad (3.1d.10)$$

so that, for v = 0,

$$J_0(z) = \frac{1}{2}\,[H_0^{(2)}(z) - H_0^{(2)}(-z)]. \qquad (3.1d.11)$$

Thus,

$$G(\vec{r},\vec{r}') = \frac{1}{2(2\pi)}\int_0^{\infty} \frac{k}{k^2-k_0^2}\,[H_0^{(2)}(k|\vec{r}-\vec{r}'|) - H_0^{(2)}(-k|\vec{r}-\vec{r}'|)]dk$$

$$= \frac{1}{4\pi}\left\{\int_0^{\infty} \frac{kH_0^{(2)}(k|\vec{r}-\vec{r}'|)}{k^2-k_0^2}dk + \int_{-\infty}^{0} \frac{kH_0^{(2)}(k|\vec{r}-\vec{r}'|)}{k^2-k_0^2}dk\right\}$$

$$= \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{k}{k^2-k_0^2}\,H_0^{(2)}(k|\vec{r}-\vec{r}'|)dk. \qquad (3.1d.12)$$

Now perform the above improper integral by the Cauchy residue formula. Assume that k_0 is replaced by $k_0 - j\epsilon$, where $\epsilon > 0$, so that

$$\{e^{-jk_0x} \leftarrow e^{-jk_0x-\epsilon x}\} \to 0 \quad \text{as} \quad x\to\infty. \qquad (3.1d.13)$$

Equation (3.1d.12) is then rewritten as

$$G(\vec{r},\vec{r}') = \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{kH_0^{(2)}(k|\vec{r}-\vec{r}'|)}{[k-(k_0-j\epsilon)][k+(k_0-j\epsilon)]}dk. \qquad (3.1d.14)$$

The half-plane of the contour is chosen so as to include the

poles that would cause G to decay to zero out at infinity. As seen from Eq. (3.1d.7), this can be guaranteed by forcing the exponent of e^{-jk_0x} to have a negative real part for positive x. This is assured, as noted in Eq. (3.1d.13), by running the contour in the lower half of the complex k-plane (see Fig. (3.1d.2)). Noting that the pole $k_0 - j\epsilon$ is included within the integration contour and $k_0 + j\epsilon$ is not, the result of integration, $-j2\pi$(sum of the residues), yields

$$G(\vec{r},\vec{r}') = \frac{-j2\pi}{4\pi} \frac{k_0}{2k_0} H_0^{(2)}(k_0|\vec{r}-\vec{r}'|) \qquad (3.1d.15a)$$

or

$$G(\vec{r},\vec{r}') = \frac{-j}{4} H_0^{(2)}(k_0|\vec{r}-\vec{r}'|) \qquad (3.1d.15b)$$

which is the form of the two-dimensional Green function that is used in this study.

3.1e Expansion of γf over a sinc basis

In the sinc basis moment method, the volume inhomogeneity function γf is expanded over a sinc basis, thereby removing the (sampled) unknown function from under the integral sign and leaving behind only, in the two-dimensional case, Hankel functions times double sinc functions. The following discussion on the expansion of a function over a sinc basis is adapted from Gabel and Roberts (1980), but all the scattered necessary ideas there are here brought concisely together, and a somewhat more general final expansion formula is obtained here than in Gabel and Roberts (1980). The sampling theorem says that a bandlimited signal

having no spectral components at or above frequency Ω can be uniquely represented by its sampled values spaced uniformly in the spatial domain at maximum spacing $1/(2(\Omega/2\pi))$. Given a function f(x), uniform sampling is mathematically equivalent to multiplying f(x) by an impulse train

$$f_s(x) = f(x) \cdot \sum_{n=-\infty}^{\infty} \delta(x-n\Delta x) \qquad (3.1e.1)$$

where Δx is the sampling interval. The function $\sum_{n=-\infty}^{\infty}\delta(x - n\Delta x)$ is periodic, so it can be expanded in a Fourier series

$$\sum_{n=-\infty}^{\infty} \delta(x-n\Delta x) = \sum_{n=-\infty}^{\infty} c_n \, e^{jn(2\pi/\Delta x)x} \qquad (3.1e.2)$$

where

$$c_n = \frac{1}{\Delta x}\int_{-\Delta x/2}^{\Delta x/2} \delta(x-n\Delta x)e^{-jn(2\pi/\Delta x)x}dx = \frac{1}{\Delta x} \qquad (3.1e.3)$$

for which $(1/\Delta x)\cdot e^{-jn2\pi x/\Delta x}$ was sampled at $n\Delta x$. Thus,

$$\sum_{n=-\infty}^{\infty} \delta(x-n\Delta x) = \frac{1}{\Delta x} \sum_{n=-\infty}^{\infty} e^{j[n(2\pi/\Delta x)]x}. \qquad (3.1e.4)$$

By multiple reversal of coordinate names in the definition of the inverse Fourier transform definition, there results the symmetry relation

$$[g(x) \Leftrightarrow \tilde{g}(k)] \Leftrightarrow [\tilde{g}(x) \Leftrightarrow 2\pi g(-k)], \qquad (3.1e.5)$$

But

$$[\delta(x) \Leftrightarrow 1] \rightarrow [1 \Leftrightarrow 2\pi\delta(k)], \qquad (3.1e.6)$$

Furthermore,

$$\mathcal{F}[g(x)e^{jk_1x}] = \int_{-\infty}^{\infty} g(x)^{-j(k-k_1)x}dx = \tilde{g}(k-k_1). \qquad (3.1e.7)$$

Thus,

$$e^{jk_1x} \Leftrightarrow 2\pi\delta(k-k_1). \tag{3.1e.8}$$

Using this transform pair, the Fourier transform of Eq. (3.1e.4) is found to be

$$\mathcal{F}[\sum_{n=-\infty}^{\infty} \delta(x-n\Delta x)] = \frac{2\pi}{\Delta x} \sum_{n=-\infty}^{\infty} \delta(k-n\frac{2\pi}{\Delta x}). \tag{3.1e.9}$$

From the frequency convolution theorem

$$g(x)h(x) \Leftrightarrow \frac{1}{2\pi} \tilde{g}(k)*\tilde{h}(k) \tag{3.1e.10}$$

the Fourier transform of $f_s(x)$ (Eq. (3.1e.1) is

$$\tilde{f}_s(k) = \frac{1}{2\pi} \tilde{f}(k)*\frac{2\pi}{\Delta x} \sum_{n=-\infty}^{\infty} \delta(k-n\frac{2\pi}{\Delta x})$$

$$= \frac{1}{\Delta x} \sum_{n=-\infty}^{\infty} \tilde{f}(k-n\frac{2\pi}{\Delta x}). \tag{3.1e.11}$$

These infinitely replicated versions of $\tilde{f}(k)$ are visualized in the diagram in Fig. (3.1e.1). If $(\pi/\Delta x) \geq \Omega$ there will be no aliasing (interference between the replications), and f(x) can be recovered by low-pass filtering $\tilde{f}_s(k)$ up to $\pi/\Delta x$. If

$$\tilde{b}_\Omega(k) = \begin{cases} 1 & |k| \leq \Omega \\ 0 & \text{otherwise} \end{cases} \tag{3.1e.12}$$

then

$$\frac{\tilde{f(k)}}{\Delta x} = \tilde{f}_s(k)\ \tilde{b}_{\frac{\pi}{\Delta x}}(k) \tag{3.1e.13}$$

so that

$$f(x) = \Delta x f_s(x) * b_{\frac{\pi}{\Delta x}}(x) \tag{3.1e.14}$$

where

$$b_{\frac{\pi}{\Delta x}}(x) = \frac{1}{2\pi}\int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{jkx}dk = \frac{1}{2\pi}\,\frac{2j\sin\,(\frac{\pi x}{\Delta x})}{jx} = \frac{1}{\Delta x}\mathrm{sinc}(\frac{x}{\Delta x}) \qquad (3.1e.15)$$

where

$$\mathrm{sinc}(\frac{x}{\Delta x}) \triangleq \frac{\sin\pi x}{\pi x}. \qquad (3.1e.16)$$

Therefore, Eq. (3.1e.14) becomes

$$f(x) = f_s(x) * \mathrm{sinc}(\frac{x}{\Delta x})$$

$$= \sum_{n=-\infty}^{\infty} f(n\Delta x)\delta(x-n\Delta x) * \mathrm{sinc}(\frac{x}{\Delta x}). \qquad (3.1e.17)$$

Finally,

$$f(x) = \sum_{n=-\infty}^{\infty} f(n\Delta x)\mathrm{sinc}[\frac{1}{\Delta x}(x-n\Delta x)]. \qquad (3.1e.18)$$

Note the difference between Eq. (3.1e.18) and Eq. (5.151) of Gabel and Roberts (1980): what is here designated as Ω replaces $\pi/\Delta x$ inside the sinc function of Eq. (3.1e.18). At first glance, it appears that the difference is due to the harmless action of low-pass filtering at Ω instead of at $\pi/\Delta x$ ($\tilde{f}(k)$ is assumed to be zero for $\Omega < k < \pi/\Delta x$), and that is what they did. However, in their case there should in general appear a scaling factor of $\Delta x\Omega/\pi$ due to using $b_\Omega(x)$ instead of $b_{\pi/\Delta x}(x)$ in Eq. (3.1e.14). That is, Eq. (3.1e.18) would in general read

$$f(x) = \Delta x(\frac{\Omega}{\pi}) \sum_{n=-\infty}^{\infty} f(n\Delta x)\mathrm{sinc}[\frac{\Omega}{\pi}(x-n\Delta x)]. \qquad (3.1e.19)$$

But note that if <u>exact</u> Nyquist sampling is used, ($\Delta x = \pi/\Omega$), the scaling factor is one. Thus, the result in Gabel and Roberts

(1980) is correct only for the case of exact Nyquist sampling. In fact, if that formula were the most general, the arguments of the sinc functions in the expansions of γf and H_0^C would not be equal because $\Omega max_{\gamma f} \neq \Omega max_{H_0^C}$, and the orthogonality of sinc functions could not be exploited. If the low-pass filter in Eq. (3.1e.12) is based on the sampling frequency rather than the maximum frequency in the signal, as was done in Eq. (3.1e.14), the result, Eq. (3.1e.18), is valid for both Nyquist sampling and oversampling.

Thus, if $\gamma(x,y)f(x,y)$ is a bandlimited function, it can be expanded over the set of sinc basis functions:

$$\gamma(x,y)f(x,y) \simeq \sum_a \sum_b \gamma(ah,bh)f(ah,bh) \cdot \left(\frac{\sin[\pi(x-ah)/h]}{\pi(x-ah)/h}\right) \left(\frac{\sin[\pi(y-bh)/h]}{\pi(y-bh)/h}\right) \qquad (3.1e.20)$$

where h is the sampling interval in both x and y directions. Substituting Eq. (3.1e.20) into Eq. (3.1c.2) results in the set of discrete, matrix equations that may be used to estimate the unknown object function $\gamma(x,y)$ at the grid points $x = ah$, $y = bh$.

A useful physical configuration (Fig. (3.1e.2)) for this problem is a set of field detectors and transmitting transducers located outside the object region, which consists of a rectangularly sampled grid completely containing the cross section of the scattering object.

Specializing Eq. (3.1c.2) with expansion of γf over the set of sinc basis functions (Eq. (3.1e.20)) to exterior (detection) points gives

$$f_{\phi m}^{sc} = \sum_j D_{mj}\gamma_j f_{\phi j} \qquad (3.1e.21)$$

and to interior (object grid) points gives

$$f_{\phi l} = f_{\phi l}^{inc} + \sum_j C_{lj}\gamma_j f_{\phi j} \qquad (3.1e.22)$$

where

ϕ = transmitter position index; $1 \leq \phi \leq n_{trans}$ where n_{trans} is the total number of transmitter positions. For each ϕ there are two independent sets of equations in Eqs. (3.1e.21) and (3.1e.22).

m = detector position index; $1 \leq m \leq n_{det}$ where n_{det} is the number of dectectors. This index m corresponds to location (x_m, y_m). For each ϕ in Eq. (3.1e.21) there are n_{det} independent equations.

l = index of location (nh,ph) in the object region;
$1 \leq l \leq N$ where N is the number of samples in the object region grid. For each ϕ in Eq. (3.1e.22) there are N independent equations. l and m correspond to (x,y) in Eq. (3.1c.2.)

j = index of location (ah,bh) in the object region;
$1 \leq j \leq N$. j corresponds to (x′,y′) in Eq. (3.1c.2).

$$D_{mj} = (-j/4) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_0^{(2)}[k_0\sqrt{(x_m-x')^2 + (y_m-y')^2}\,] \cdot \left(\frac{\sin[\pi(x'-ah)/h]}{\pi(x'-ah)/h}\right)\left(\frac{\sin[\pi(y'-bh)/h]}{\pi(y'-bh)/h}\right) dx'dy' \qquad (3.1e.23)$$

$$C_{lj} = (-j/4) \int\int_{-\infty}^{\infty} H_0^{(2)}[k_0\sqrt{(nh-x')^2 + (ph-y')^2}\,] \cdot \left(\frac{\sin[\pi(x'-ah)/h]}{\pi(x'-ah)/h}\right) \left(\frac{\sin[\pi(y'-bh)/h]}{\pi(y'-bh)/h}\right) dx'dy'. \quad (3.1e.24)$$

Many useful details for the computation of D_{mj} and C_{lj} can be found in (Tracy and Johnson, 1983), and are made more explicit and complete in the section on computational results below.

3.1f Note concerning validity of the internal field equations

It is stated in (Cavicchi, 1986) that within the cylinder, the incident field does not exist because it does not satisfy the wave equation there ($k_{cyl} \neq k_0$ so $(\nabla^2 + k_{cyl}^2) f^{inc} \neq 0$). One might, then, question the validity of Eq. (3.1e.22):

$$f_{\phi l} = f_{\phi l}^{inc} + \sum_j C_{lj}\gamma_j f_{\phi j} \quad (3.1f.1)$$

evaluated <u>within</u> the object region, for the same reason. The controversy is resolved, however, when one realizes the assumptions behind Eq. (3.1e.22). In Eq. (3.1e.22) a distribution of point sources is assumed, each radiating into <u>free space</u>. The continuous space form of Eq. (3.1e.22) is Eq. (3.1c.2):

$$f(x,y) = f^{inc}(x,y) - \frac{j}{4} \int\int_{-\infty}^{\infty} \gamma(x',y') f(x',y') H_0^{(2)}[k_0\sqrt{(x-x')^2 + (y-y')^2}\,] dx'dy' \quad (3.1f.2)$$

where $H_0^{(2)}(k_0\sqrt{\cdot})$ is the <u>free-space</u> (unbounded region) Green function. Thus, each point source, of complex strength $\gamma(x',y')f(x',y')$, radiates a wave that propagates away from the

source as if it were in isolation. However, this is not assuming single scattering only, for the induced "source" strength $\gamma(x',y')f(x',y')$ includes all diffraction effects by virtue of $f(x',y')$ being the exact total field due to all interactions of the incident field with the inhomogeneous medium. A superposition of solutions of varying strengths $A(\vec{r}')$ of

$$(\nabla^2 + k_0^2) f_i(\vec{r}) = A(\vec{r}')\delta(\vec{r}-\vec{r}') \qquad (3.1f.3)$$

is used, where the strengths $A(\vec{r}')$ account for all scattering effects.

This view is different from, but mathematically equivalent to the one used in calculating the exact total field within a scattering cylinder. There, instead of numerous equivalent point sources radiating into free space, one speaks of a source-free region with a propagation constant different from k_0 (i.e., k_{cyl}), and in the general inhomogeneous region, $k(\vec{r})$. But in the case of the cylinder, because of its fundamental shape in cylindrical coordinates, one can simply match expressions for the boundary conditions of continuous pressure and radial velocity at the cylinder surface and also have, in particular, the given expressions for the field within the (homogeneous) cylinder which satisfy the wave equation valid there:

$$(\nabla^2 + k_{cyl}^2) f = 0. \qquad (3.1f.4)$$

3.1g Alternating variables method of moments

In terms of the standard method of moments terminology, the

basis functions are shifted sinc functions with shift argument being the location of a point in the object region in both Eqs. (3.1e.21) and (3.1e.22). The testing functions are Dirac delta functions (point matching), where the shift argument of the delta functions is in Eqs. (3.1e.21) the location of a detector and in Eqs. (3.1e.22) the location of a point in the object region. In fact, point matching is equivalent to the discretization of the continuous integral equation in Eq. (3.1c.2).

Note that in Eqs. (3.1e.21) and (3.1e.22) there are two unknowns: f and γ. Following Johnson and Tracy's (1983) algorithm a single iteration comprises the three steps: (a) Initial guesses ($\tilde{f}^0 = f^{inc}$, $\tilde{\gamma}^0 = 0$) are assigned. (Here $\tilde{\gamma}$ represents an estimate of γ, not its Fourier transform, as in Chapter 2. In subsequent uses, the meaning of the symbol should be clear from the context.) (b) Eq. (3.1e.21) is solved for $\tilde{\gamma}^1$: by way of the Algebraic Reconstruction Technique (ART) using scattered field data obtained as described below, holding $f = \tilde{f}^0$ fixed. A version of ART was developed using a paper by Herman et. al. (1973). (A significant difference from previous uses of ART is that here the relevant matrices are nonsparse.) Note that Eqs. (3.1e.21) become a set of linear equations for $\tilde{\gamma}^1$. (c) Now holding $\tilde{\gamma}^1$ fixed, Eqs. (3.1e.22) are solved for $\tilde{f}^1$ for all ϕ. Steps (b) and (c) are repeated until convergence is obtained.

3.1h Geometrical argument for the algebraic reconstruction technique (ART, or Kaczmarz method)

The following is a geometrical/algebraic argument that the Algebraic Reconstruction Technique correction is indeed a means to obtain the orthogonal projection of the current estimate of the unknown vector onto the hyperplane represented by one row vector of the row space of the given matrix. Here the case of a two-dimensional matrix problem is examined, which can easily be generalized to an n-dimensional problem. Given the following 2x2 matrix problem:

$$R\vec{r}_a = \begin{pmatrix} A & B \\ D & E \end{pmatrix} \begin{pmatrix} x_a \\ y_a \end{pmatrix} = \begin{pmatrix} C \\ F \end{pmatrix} = \vec{G}, \qquad (3.1h.1)$$

the first row of the matrix represents the vector $\vec{1}$ for which the inner product with (x_a,y_a) equals C:

$$(A,B)\cdot(x_a,y_a) = Ax_a + By_a = C. \qquad (3.1h.2)$$

Here, $\vec{r}_a = (x_a,y_a)$ is the unknown vector and C is one element of the known (e.g., measured) vector $\vec{G}$. In Fig. (3.1h.1) is drawn the line Ax + By = C, the general form of the first equation contained in the matrix equation Eq. (3.1h.1). Also shown is an initial guess for the solution of Eq. (3.1h.1), $\vec{r}_0 = (x_0,y_0)$, and the true solution vector $\vec{r}_a = (x_a,y_a)$, which of course is also a point on the line, and thus satisfies $Ax_a + By_a = C$. (The vector $\vec{r}_a$ also satisfies $Dx_a + Ey_a = F$, but here only the first row of Eq. (3.1h.1) is being considered.) The distance t is the distance from $\vec{r}_0$ to the line, along the normal to the line and passing

through $\vec{r}_0$. The x and y intercepts of Ax + By = C are found by setting y and x equal to zero, respectively, in Ax + By = C. If C = 0, both intercepts are zero; the slope of the line is then easily found to be -A/B. Therefore, the slope of the perpendicular is B/A. Thus, the unit vector normal to the line Ax + By = C is

$$\hat{n} = \frac{(A,B)}{\sqrt{A^2+B^2}} . \qquad (3.1h.3)$$

Next, t is found, the component of the error vector $\vec{l} = (x_0 - x_a, y_0 - y_a)$ normal to the line Ax + By = C. So

$$t = \vec{l} \cdot \hat{n}$$

$$= \frac{A(x_0-x_a)+B(y_0-y_a)}{\sqrt{A^2 + B^2}}$$

$$= \frac{Ax_0+By_0-(Ax_a+By_a)}{\sqrt{A^2 + B^2}}$$

$$= \frac{Ax_0+By_0-C}{\sqrt{A^2 + B^2}} . \qquad (3.1h.4)$$

(Thus, only C is necessary, typically a measurement of the scattered or incident field in tomography, and not $\vec{r}_a$, the desired solution.) Therefore, to "move" the initial guess to its orthogonal projection onto the line Ax + By = C, one simply adds $-t\hat{n}$ to $\vec{r}_0$. This correction can be written

$$\vec{r}_1 = \vec{r}_0 + \left[\frac{C-(Ax_0+Bx_0)}{A^2 + B^2}\right](A,B) \qquad (3.1h.5)$$

or

$$\vec{r}_i = \vec{r}_{i-1} + \left(\frac{G_i - \vec{R}_i \cdot \vec{r}_{i-1}}{|\vec{R}_i|^2}\right)\vec{R}_i \qquad (3.1h.6)$$

for a general correction using row i, $\vec{R}_i$, to correct guess i -1, $\vec{r}_{i-1}$, in the matrix equation $R\vec{r} = \vec{G}$. An example of repetitive cycling through the rows of the matrix equation using the above correction on the initial guess is shown for the two-dimensional case in Fig. (3.1h.2); its effectiveness there is quite convincing. The Algebraic Reconstruction Technique correction as used in, for example, computer tomography, is an obvious generalization of the above correction to higher dimensions (where, for example, the lines discussed here become hyperplanes).

A short discussion of this generalization follows. The definition of a (hyper)plane is all $\vec{r}$ such that, for a specified normal (unit) vector $\hat{n}$ and a given point $\vec{r}_a$ in the plane, $\hat{n}\cdot(\vec{r} - \vec{r}_a) = 0$. Written in the form of the row of a matrix, $\hat{n}\cdot\vec{r} = \hat{n}\cdot\vec{r}_a$, the components can be identified with a row of $R\vec{r}_a = \vec{G}$: $\hat{n} = \vec{R}_i/||\vec{R}_i||$ and $\hat{n}\cdot\vec{r}_a = G_i/||\vec{R}_i||$ for row i. Thus, the $\vec{R}_i$ specifies the normal to the hyperplane and G_i is the inner product that the position vector of any point in the plane must have with the normal $\vec{R}_i$. (A distinction should be made between points in a hyperplane specified by position vectors and vectors actually contained within a hyperplane. Position vectors are generally specified with respect to the zero vector, but a hyperplane is specified by forcing the orthogonality of vectors emanating from the point $\vec{r}_a$ with a normal passing through the point $\vec{r}_a$ on the plane. Hence, for points $\vec{r}$ on the hyperplane,

$\vec{r}.\hat{n} = G_i/||\vec{R}_i|| \neq 0.$) So the component of the error vector $\vec{l}$ $= \vec{r}_0 - \vec{r}_a$ of the initial guess $\vec{r}_0$ normal to the line vector $\vec{R}_i$

$$t = \vec{l}.\hat{n}$$
$$= (\vec{r}_0 - \vec{r}_a)\cdot\vec{R}_i/||\vec{R}_i||$$
$$= (\vec{r}_0\cdot\vec{R}_i - G_i)/||\vec{R}_i|| \qquad (3.1h.7)$$

Multiplying t by $\hat{n}$ and subtracting the result from $\vec{r}_0$, Eq. (3.1h.6) follows directly.

3.1i Mathematical character of the ART corrections

Possibly the best figure of merit for a given estimation of γ and f in the absence of knowledge of γ^{ex} is the mean-squared error in the resulting estimation of the measured scattered field, through Eq. (3.1e.21)

$$E(\gamma) = (f_\phi^{sc} - \sum_j D_{mj}\gamma_j f_{\phi j})^T (f_\phi^{sc} - \sum_j D_{mj}\gamma_j f_{\phi j}) \qquad (3.1i.1)$$

where here $\sum_j D_{mj}\vec{\gamma}_j\vec{f}_{\phi j}$ and $\vec{f}_\phi^{sc}$ are vectors of length $n_{trans}n_{det}$ and $\vec{\gamma}$ and $\vec{f}_\phi$ are vectors of length n_{max}^2. The solution for $\vec{\gamma}$ to minimize E is seen to be a specific case of the solution for $\vec{x}$ in the generic matrix equation $A\vec{x} = \vec{y}$ minimizing

$$E(\vec{x}) = (\vec{y}-A\vec{x})^T(\vec{y}-A\vec{x}). \qquad (3.1i.2)$$

Let $\vec{x}^*$ be $\vec{x}$ such that $E(\vec{x})$ is minimized. Then

$$A^T(\vec{y}-A\vec{x}^*) = 0. \qquad (3.1i.3)$$

Two proofs will be given, similar to those given in (Papoulis, 1984) but here given for the matrix rather than the scalar case,

and in the deterministic rather than the stochastic setting. To minimize $E(\vec{x})$, set

$$\frac{\partial E(\vec{x})}{\partial \vec{x}} (\vec{x}^*) = 0 \qquad (3.1i.4)$$

where the derivative of a scalar function ϕ of a vector $\vec{x}$ with respect to the vector $\vec{x}$ equals the vector

$$\frac{\partial \phi}{\partial \vec{x}} = \frac{\partial \phi}{\partial x_1}\hat{i} + \frac{\partial \phi}{\partial x_2}\hat{j} + \frac{\partial \phi}{\partial x_3}\hat{k} \quad . \; . \; . \qquad (3.1i.5)$$

Expanding Eq. (3.1i.2),

$$\begin{aligned} E(\vec{x}) &= \vec{y}^T\vec{y} - \vec{y}^TA\vec{x} - (A\vec{x})^T\vec{y} + (A\vec{x})^TA\vec{x} \\ &= \vec{y}^T\vec{y} - \vec{y}^TA\vec{x} - \vec{x}^TA^T\vec{y} + \vec{x}^TA^TA\vec{x}. \end{aligned} \qquad (3.1i.6)$$

Differentiating $E(\vec{x})$ with respect to $\vec{x}$ and setting the result equal to zero,

$$\frac{\partial E(\vec{x})}{\partial \vec{x}} (\vec{x}^*) = -2\vec{y}^TA + 2\vec{x}^{*T}A^TA = 0. \qquad (3.1i.7)$$

Rearranging,

$$A^T(\vec{y}-A\vec{x}^*) = 0. \qquad (3.1i.8)$$

The second proof calculates $E(\vec{x}^b)$, where $\vec{x}^b \neq \vec{x}^*$, and shows that it is greater than or equal to $E(\vec{x}^*)$:

$$\begin{aligned} E(\vec{x}^b) &= (\vec{y}-A\vec{x}^b)^T(\vec{y}-A\vec{x}^b) \\ &= [\vec{y}-A\vec{x}^*+A(\vec{x}^*-\vec{x}^b)]^T[\vec{y}-A\vec{x}^*+A(\vec{x}^*-\vec{x}^b)] \\ &= (\vec{y}-A\vec{x}^*)^T(\vec{y}-A\vec{x}^*) + (\vec{x}^*-\vec{x}^b)^TA^TA(\vec{x}^*-\vec{x}^b) \\ &\quad + (\vec{y}-A\vec{x}^*)^TA(\vec{x}^*-\vec{x}^b) + [A(\vec{x}^*-\vec{x}^b)]^T(\vec{y}-A\vec{x}^*) \end{aligned} \qquad (3.1i.9)$$

where the second term is a scalar of the form $\vec{x}^TW\vec{x}$, where W is a positive definite symmetric matrix. This term is therefore a quadratic term, always nonnegative. The third and fourth terms

are zero by the definition of $\vec{x}^*$. (The third term equals the fourth term because it, being a scalar, equals its transpose, the fourth term.) Consequently,

$$E(\vec{x}^b) \geq (\vec{y}-A\vec{x}^*)^T(\vec{y}-A\vec{x}^*) = E(\vec{x}^*). \qquad (3.1i.10)$$

Thus, $\vec{x}^*$ is an optimum solution for the minimization of $E(\vec{x})$.

Attention is now turned to the problem of solving Eq. (3.1i.8) for $\vec{x}^*$. It is obvious that

$$A^TA\vec{x}^* = A^T\vec{y} \qquad (3.1i.11)$$

so it will be necessary to invert A^TA, which may be as difficult to solve as the original matrix inversion problem $A\vec{x} = \vec{y}$, especially with respect to storage and efficiency, but also in the inability to use prior information about $\vec{x}^*$. In any case this is the pseudoinverse problem:

$$\vec{x}^* = (A^TA)^{-1}A^T\vec{y} \qquad (3.1i.12)$$

where $(A^TA)^{-1}A^T$ is the pseudoinverse of A. But, according to the Algebraic Reconstruction Technique, given an initial guess $\vec{x}^i$, it can be moved successively closer to $\vec{x}^*$ by the correction

$$\vec{x}^{i+1} = \vec{x}^i + \beta\frac{(y_i-\vec{a}_i\cdot\vec{x}^i)}{||\vec{a}_i||^2}\vec{a}_i^T \qquad (3.1i.13)$$

where $\vec{x}^i$ is a column vector, $\vec{a}_i$ is a row (vector) of A written in the form of a column vector--that is, $y_i = \vec{a}_i\cdot\vec{x}$, and β is a relaxation constant. By doing this $\vec{x}^{i+1}$ is forced to be such that

$$\vec{a}_i^T\{\vec{y}_i-\vec{a}_i\cdot[\vec{x}^i + \beta\frac{(y_i-\vec{a}_i.\vec{x}^i)}{||\vec{a}_i||^2}\vec{a}_i^T]\}$$

$$= \vec{a}_i^T[y_i - \vec{a}_i.\vec{x}_i - \frac{\vec{a}_i.y_i\vec{a}_i^T}{||\vec{a}_i||^2} + \frac{\vec{a}_i.(\vec{a}_i.\vec{x}^i)}{||\vec{a}_i||^2}\vec{a}_i^T]$$

$$= 0 \qquad (3.1i.14)$$

for $\beta = 1.0$. From the scalar/vector form of Eq. (3.1i.8) ($\vec{y}$ a scalar and A and $\vec{x}$ vectors) and the proofs given above, it is evident that, a row at a time, corrections are made on $\vec{x}^i$ which obtain the least squares solution of $y_i = \vec{a}_i.\vec{x}$ (for a unity relaxation constant β). From another viewpoint, the correction removes the component of $\vec{x}_i$ that is orthogonal to the hyperplane i, the intersection of all hyperplanes being $\vec{x}^*$ for a well-conditioned full rank problem. The following statements can be proved (Tanabe, 1971). If the matrix has full rank $\vec{x}^i$ converges to the true solution for i -> ∞. For underdetermined systems it will converge to the local least squares solution closest to $\vec{x}^0$, the initial guess. For overdetermined inconsistent systems, the solution will end up in limit cycles. However, these statements are strictly valid only for the case of using unity relaxation constants. In the use of ART in this thesis, a relaxation constant as low as 0.2 was used (see Section 3.3i). [Note that use of $\beta < 1$ moves the solution estimate less for a particular row, and therefore gives less emphasis to that measurement. One could, by using lower β for less confident measurements, incorporate such knowledge into the iterative corrections. In this thesis, however, the same value of β was used

for all matrix rows (measurements). Conversely, use of $\beta > 1$ would always be unstable because of the opposite effect it has compared with $\beta < 1$.] Furthermore, for weak scattering problems, accuracy of the reconstructions stayed high even in the presence of over 10% noise added to the scattered field data, the noise making the equations inconsistent. Table (3.1i.1) illustrates this fact for a ka = 12.6, 5% speed of sound mismatch, lossless cylinder. (All numbered tables can be found at the end of this thesis.) The squared error in $\tilde{\gamma}$ is given after iterations 1 and 7; significant degradation (increase in squared error relative to zero noise) occurs between 10% and 20% additive noise to the scattered field for this particular weak scattering problem. For each level of noise considered in Table (3.1i.1), the inner product of all intermediate estimations (resulting from orthogonal projections on each of the hyperplanes in the matrix equation) of γ with that at the beginning of the iteration were calculated and examined during iteration 8, by which time the reconstructions as usually presented (before row correction 1) had converged. For example, for the case of zero noise, the squared error in γ was 253.2 after seven iterations and 254.8 after eight iterations. These inner products changed from row to row only very slightly and gradually. For example, the inner product magnitude might gradually reduce from 1.0 to 0.997 for row 900 and later wander back to 0.9999 by the end of the iteration (these values actually occurred for the case of 20% noise; for lower noise, the inner products changed even less). Thus, the magnitude of limit cycles is computationally negligible, at least for the case of weak

scattering. As a side note, it can be shown (Brogan, (1982)) that in the least squares case (overdetermined/inaccurate system) the pseudoinverse solution of Eq. (3.1i.8) is the unique $\bar{x}$ of minimum norm that specifies the coefficients of the expansion of the projection of $\bar{y}$ onto the row space of A (which corresponds to the unique solution generated by the pseudoinverse matrix $(A^TA)^{-1}A^T$, which uniquely satisfies the four Moore-Penrose conditions (Golub and VanLoan, 1984).

A final observation here would be that in cases where the incident field is sufficiently inaccurate for representing the total internal field, the solution to which the present algorithm converges may no longer be the desired solution. Complicating matters is the error in both scattered field and internal field equations due to the discretization of the integral equations, which will be shown to be significant when using QR decomposition for a least squares solution (see Section 5.3a). The fact that relaxation constants as low as 0.2 have been necessary may indicate how poorly correcting for one measurement at a time can simultaneously correct the estimate for the entire matrix problem (that is, for the desired solution). Further investigations could analyze the solutions obtained using unity relaxation constants, for example, comparing them with the full matrix least squares solution. However, for numerous numerical studies the resulting solution vector converges in one or two passes through the equations to one acceptably close to the desired solution for small to moderately large normed exact (chosen) solutions.

3.2 Programming Details

3.2a Incident field

For an incident field a cylindrical line source is used (proportional to the two-dimensional free-space Green function):

$$f_{\phi j}^{inc} = H_0^{(2)}[k_0\sqrt{(x_\phi - ah)^2 + (y_\phi - bh)^2}\,] \qquad (3.2a.1)$$

where (ah,bh) = jth pixel coordinates vector. At all points at which it is to be evaluated, this incident field satisfies the homogeneous Helmholtz wave equation, as it must. In this implementation a circular ring of transmit/receive transducers is centered on the center of a square discretized object region. Therefore, the object region point-transducer distances are found by the law of cosines (see Fig. (3.1e.2)). The real and imaginary parts of the incident field, extending from the source over the object region (boxed in) are illustrated in Fig. (3.2a.1).

3.2b Pixel coordinates

A programming relation used repeatedly is that between pixel number and rectangular location. For example, the rectangular location of pixel j is

$$(a,b) = [\text{int}(\frac{j-1}{n_{max}}) + 1,\ (j-1)_{\text{mod}(n_{max})} + 1] \qquad (3.2b.1)$$

where mod denotes modulus and n_{max} is the number of pixels on a side of the object region: $n_{max}^2 = N$. Going the other way,

$$j = (a-1)n_{max} + b. \qquad (3.2b.2)$$

3.2c Coefficient generation

The integrals in Eqs. (3.1e.23) and (3.1e.24) are difficult to perform numerically because of the infinite domain of integration, combined with the singularity of $H_0^{(2)}(\cdot)$ at the origin. The goal is to expand $H_0^{(2)}(\cdot)$ over a sinc basis (see Section 3.1e) and then use the orthogonality property of sinc functions (see Section 2.3a) to evaluate the resulting integrals of quadruple sinc functions. The integrals for C_{lj} and D_{mj} are specializations of the form

$$W_j(x,y) = -\frac{j}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\sin[\frac{\pi}{h}(x'-ah)]}{\frac{\pi}{h}(x'-ah)} \frac{\sin[\frac{\pi}{h}(y'-bh)]}{\frac{\pi}{h}(y'-bh)} \cdot H_0^{(2)}[k_0\sqrt{(x-x')^2 + (y-y')^2}]dx'dy' \qquad (3.2c.1)$$

The problem arises because for the sample at the origin, the imaginary part of $H_0^{(2)}(\cdot)$, $-Y_0(\cdot)$, is unbounded. But, the sinc function expansion is valid only for finite functions (McNamee, Stenger et al., 1971). A successful treatment (Tracy and Johnson, 1983) of this problem has been to separate the integration into two parts, one for the distance between (x,y) and (x′,y′) less than some small quantity, x_z, for which the integration is carried out by two-dimensional numerical quadrature which allows for singularities, and one for distances larger than the small quantity. This second integral can be thought of as the integral over the entire domain of a modified function that is nonsingular at the origin, and so can without difficulty be expanded over a

sinc basis as desired. This section reviews the details of this method.

Let $H_0^C(\cdot)$ and $Y_0^E(\cdot)$ be defined as follows:

$$H_0^C(x) = \begin{cases} H_0^{(2)}(x) & x \geq x_z \\ \text{Function with no singularity at } x=0 & x < x_z \end{cases} \tag{3.2c.2}$$

$$Y_0^E(x) = \begin{cases} 0 & x \geq x_z \\ \text{Error caused by using } H_0^C(x) & x < x_z \end{cases} \tag{3.2c.3}$$

so that

$$H_0^{(2)}(x) = H_0^C(x) - jY_0^E(x). \tag{3.2c.4}$$

To be more specific, introduce, as do Tracy and Johnson (1983), the "capping function" $F_C(x)$ as follows. First note that

$$H_0^{(2)}(x) = J_0(x) - jY_0(x) \tag{3.2c.5}$$

where $Y_0(\cdot)$ is singular at the origin. Define the function $Y_0^C(x)$ as

$$Y_0^C(x) = \begin{cases} Y_0(x) & x \geq x_z \\ F_C(x) & x < x_z \end{cases} \tag{3.2.6}$$

where x_z is the (for example) second zero of $Y_0(\cdot)$ and $F_C(\cdot)$ is any smooth function satisfying

$$\begin{aligned} F_C(x_z) &= Y_0(x_z) \quad (=0) \\ F'_C(x_z) &= Y'_0(x_z) \\ F'_C(0) &= 0. \end{aligned} \tag{3.2c.7}$$

A convenient function to choose for F_C is a sinc function:

$$F_C(x) = A\frac{\sin\left(\pi\frac{x}{x_z}\right)}{\pi\frac{x}{x_z}} \qquad (3.2c.8)$$

where A is a constant chosen so that the derivative of $F_C(\cdot)$ matches that of $Y_0(\cdot)$ at $x = x_z$. To obtain the derivative $Y_0(x_z)$ two methods can be used. One is to divide the difference between two values of $Y_0(\cdot)$ just on either side of x_z by the differences in x values. That is, given that $x_z = 3.9576784$,

$$Y'_0(x_z) \simeq \frac{Y_0(3.96) - Y_0(3.94)}{3.96-3.94} \simeq \frac{-0.0009343-0.0071319}{3.96-3.94}$$

$$= -0.40331. \qquad (3.2c.9)$$

The other method is to use the relation

$$xY'_n(x) - nY_n(x) = -xY_{n+1}(x) \qquad (3.2c.10)$$

which, for n = 0 reduces to

$$Y'_0(x) = -Y_1(x). \qquad (3.2c.11)$$

The value $Y_1(x_z)$ equals 0.402543, which is very close to that obtained for the first method. Now compute $F_C'(x_z)$ as follows:

$$F'_C(x) = A\{\frac{\frac{\pi x}{x_z}\left(\frac{\pi}{x_z}\right)\cos\left(\frac{\pi x}{x_z}\right) - \left(\frac{\pi}{x_z}\right)\sin\left(\frac{\pi x}{x_z}\right)}{(\pi\frac{x}{x_z})^2}\} = A\{\frac{\left(\frac{\pi x}{x_z}\right)\cos\left(\frac{\pi x}{x_z}\right) - \sin\left(\frac{\pi x}{x_z}\right)}{\left(\frac{\pi x}{x_z}\right)^2}\}$$

$$F'_C(x_z) = A\{\frac{-\pi-0}{\pi x_z}\} = \frac{-A}{x_z} = -0.4025429 \qquad (3.2c.12)$$

so that $A = 0.4025429(3.9576784) = 1.5931353$ and therefore

$$F_C(x) = 1.5931353\ \{\frac{\sin\left(\pi\frac{x}{3.9576784}\right)}{\left(\pi\frac{x}{3.9576784}\right)}\} = 2.0069812\frac{\sin(0.7937969x)}{x}. \qquad (3.2c.13)$$

Thus, a sketch of $F_C(\cdot)$ and $Y_0(\cdot)$ appears as in Fig. (3.2c.1). As an aside, note that $F_C'(0)$ is

$$F'_C(0) = \frac{"0"}{0} = A\left.\frac{\frac{d}{dx}\{x\frac{\pi}{x_z}\cos\left(\frac{\pi x}{x_z}\right) - \sin\left(\frac{\pi x}{x_z}\right)\}}{\frac{d}{dx}\{\pi\frac{x^2}{x_z}\}}\right|_{x\to 0}$$

$$= A\left.\frac{x(\frac{\pi}{x_z})^2\cdot\sin\left(\frac{\pi x}{x_z}\right) + \frac{\pi}{x_z}\cos\left(\frac{\pi x}{x_z}\right) - \frac{\pi}{x_z}\cos\left(\frac{\pi x}{x_z}\right)}{2\frac{\pi x}{x_z}}\right|_{x\to 0} = \frac{"0"}{0}\ \text{again}$$

$$= A\left.\frac{\frac{d}{dx}\{(\frac{\pi}{x_z})^2 x\sin\left(\frac{\pi x}{x_z}\right)\}}{\frac{d}{dx}\{2\frac{\pi x}{x_z}\}} = \frac{(\frac{\pi}{x_z})^2\sin\left(\frac{\pi x}{x_z}\right) + (\frac{\pi}{x_z})^3 x\cos\left(\frac{\pi x}{x_z}\right)}{\frac{2\pi}{x_z}}\right|_{x\to 0} = 0 \qquad (3.2c.14)$$

as expected. The error caused by using $Y_0{}^C(x)$ is

$$Y_0{}^E(x) = \begin{cases} Y_0(x) - F_c(x) & x<x_z \\ 0 & x\geq x_z \end{cases} \qquad (3.2c.15)$$

so

$$H_0{}^{(2)}(x) = J_0(x) - jY_0{}^C(x) - jY_0{}^E(x) \qquad (3.2c.16a)$$

or

$$H_0^{(2)}(x) = H_0^C(x) - jY_0^E(x). \tag{3.2c.16b}$$

The result of substituting this expansion into Eq. (3.2c.1) is

$$W_j(x,y) = -\frac{j}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_0^C(k_0\sigma)\frac{\sin[\frac{\pi}{h}(x'-ah)]}{\frac{\pi}{h}(x'-ah)}\frac{\sin[\frac{\pi}{h}(y'-bh)]dx'dy'}{\frac{\pi}{h}(y'-bh)}$$

$$-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Y_0^E(k_0\sigma)\frac{\sin[\frac{\pi}{h}(x'-ah)]}{\frac{\pi}{h}(x'-ah)}\frac{\sin[\frac{\pi}{h}(y'-bh)]}{\frac{\pi}{h}(y'-bh)}\,dx'dy' \tag{3.2c.17}$$

where

$$\sigma = \sqrt{(x-x')^2 + (y-y')^2} \tag{3.2c.18}$$

and where the first integral is over all space, while the second has contributions only out to $\sigma = x_z/k_0$. Now $H_0^C(\cdot)$ is almost bandlimited to a little over k_0 (see Fig. (3.2e.1) for the real part of the Fourier transform of $H_0^{(2)}$), so it can be expanded over a sinc basis. The resulting integration over that term is then trivial, because of the orthogonality of sinc functions (see Section 2.3a), so that one is left with a much smaller area of integration over only the term $Y_0^E(x)$.

Expand $H_0^C(\cdot)$ over a sinc basis:

$$H_0^C[k_0\sqrt{(x-x')^2 + (y-y')^2}] \simeq \sum_c \sum_d H_0^C[k_0\sqrt{(x-ch)^2 + (y-dh)^2}]$$

$$\cdot\{\frac{\sin[\frac{\pi}{h}(x'-ch)]}{\frac{\pi}{h}(x'-ch)}\}\{\frac{\sin[\frac{\pi}{h}(y'-dh)]}{\frac{\pi}{h}(y'-dh)}\}. \tag{3.2c.19}$$

The first integration can consequently be written as

$$-\frac{j}{4}\sum_c\sum_d H_0^C[k_0\sqrt{(x-ch)^2+(y-dh)^2}]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{\frac{\sin[\frac{\pi}{h}(x'-ch)]}{\frac{\pi}{h}(x'-ch)}\}$$

$$\cdot\{\frac{\sin[\frac{\pi}{h}(y'-dh)]}{\frac{\pi}{h}(y'-dh)}\}\{\frac{\sin[\frac{\pi}{h}(x'-ah)]}{\frac{\pi}{h}(x'-ah)}\}\{\frac{\sin[\frac{\pi}{h}(y'-bh)]}{\frac{\pi}{h}(y'-bh)}\}\,dx'dy' \qquad (3.2c.20)$$

Because the only nonzero term occurs for c = a and d = b, the result is

$$-\frac{j}{4}h^2H_0^C[k_0\sqrt{(x-ah)^2+(y-bh)^2}]. \qquad (3.2c.21)$$

Now only the second term is left:

$$-\frac{j}{4}(-j)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}Y_0^E(k_0\sigma)\{\frac{\sin[\frac{\pi}{h}(x'-ah)]}{\frac{\pi}{h}(x'-ah)}\}\{\frac{\sin[\frac{\pi}{h}(y'-bh)]}{\frac{\pi}{h}(y'-bh)}\}dx'dy' \qquad (3.2c.22)$$

The first step is to convert fully to polar form, as illustrated in Fig. (3.2c.2). Call the result R(x,y,ah,bh). The coordinate transformation is

$$x' = x + \sigma\cos\phi$$

$$y' = y + \sigma\sin\phi. \qquad (3.2c.23)$$

Depending on location of field evaluation (internal field point or remote detector) the sum of the results of the two integrations in Eq. (3.2c.17) are called C_{1j} and D_{mj}, respectively. The basic geometries for these two sets of coefficients are shown in Fig. (3.2c.3).

3.2c1 Internal field equations coefficients C_{1j}

First C_{1j} will be discussed. In this case the result of the first integration is simply written

$$h^2 H_0{}^C[k_0 h\sqrt{(n-a)^2 + (m-b)^2}] \qquad (3.2c1.1)$$

where the observation point is (nh,mh). The other integration is R(nh,mh,ah,bh). In polar form, this integral is written

$$R(ah,bh,nh,mh) = -\frac{j}{4}(-j)\int_0^{2\pi}\int_0^{\left(\frac{x}{z}\right)k_0} Y_0{}^E(k_0\sigma)\left[\frac{\sin\pi(n-a+\frac{\sigma}{h}\cos\phi)}{\pi(n-a+\frac{\sigma}{h}\cos\phi)}\right]$$

$$\cdot\left[\frac{\sin\pi(m-b+\frac{\sigma}{h}\sin\phi)}{\pi(m-b+\frac{\sigma}{h}\sin\phi)}\right]\sigma d\sigma d\phi. \qquad (3.2c1.2)$$

These integrals can be evaluated, as noted above, by two-dimensional numerical quadrature routines that can handle a singularity at the origin.

3.2c1a Two-dimensional quadrature

Because no two-dimensional quadrature routine allowing singularities was available, some minor modifications were made to a one-dimensional quadrature routine which was subsequently placed inside another copy of itself. To check out the new two-dimensional routine, the finite radius two-dimensional integral of the natural log was chosen. The one-dimensional integral is

$$I(a) = \int_0^a \ln x\,dx = x\ln x - x\Big]_a^0 = a\ln a - a - [0\ln 0 - 0]$$

$$= a(\ln a - 1) - \left[\left(\frac{\lim_{x\to 0} x}{\lim_{x\to 0}\frac{1}{x}} = \lim_{x\to 0}\frac{\frac{1}{x}}{\frac{1}{x^2}} = 0\right) - 0\right]$$

$$= a(\ln a - 1). \qquad (3.2c1a.1)$$

As a test of correctness of the routine, the following specific

cases were investigated: I(1) = ln(1) - 1 = -1 and I(2) = 2(ln(2) - 1) = -0.613706. The two-dimensional integration is

$$I_1(a) = \int_0^a \int_0^a \ln x \ln y \, dx dy = I^2(a) \qquad (3.2c1a.2)$$

and using (-a,a) as the integration interval to embed the singularity within the integration region gives

$$I_2(a) = \int_{-a}^{a} \int_{-a}^{a} \ln|x| \ln|y| \, dx dy = 4I^2(a). \qquad (3.2c1a.3)$$

Numerically, the results came within the tolerance as specified in the calling statement. Because the test cases worked, one could have high confidence that for other similar singular integrations, reliable results could be obtained; in the sinc basis method, these are zeroth order Bessel Functions of the second kind times double shifted sinc function. Comparing the singular behavior of $Y_0(x)$ with that of ln(x),

x	ln(x)	$Y_0(x)$
0.02	-3.91	-2.6
0.0001	-9.2	-5.90

illustrates numerically that $Y_0(x)$, the singular function that must be numerically integrated, in being less steep than ln(x) near the origin, is no worse than a logarithmic singularity, and consequently one can have high confidence in the accuracy of the numerically calculated term of the matrix coefficients.

3.2c1b Computational characteristics of the coefficients C_{lj}

In this section the various elements making up the numerical integration term of the C_{lj} coefficients will be explored pictorially. First, the shifted sinc function will be defined and investigated. The polar coordinates form of the shifted double sinc function, used in the numerical quadrature, is

$$S_{n,m;a,b}(x,y) = \text{sinc}(n-a+\tfrac{\sigma}{h}\cos\phi)\ \text{sinc}(m-b+\tfrac{\sigma}{h}\sin\phi) \qquad (3.2c1b.1)$$

where

$$\sigma = h\sqrt{(\tfrac{x}{h} - n)^2 + (\tfrac{x}{h} - m)^2} \qquad (3.2c1b.2)$$

and where

$$\phi = \tan^{-1}(\tfrac{y-mh}{x-nh}). \qquad (3.2c1b.3)$$

(Remember that the generic form (x,y) is used here, but for purposes of integration the axes would be labeled (x′,y′).) The pixel of summation in Eq. (3.1e.22) is j = (a,b) and the pixel of evaluation there is l = (n,m). Although S appears to depend on (n,m) it does not; the cosϕ and n in the argument cancel out the dependence on n in σ and then in the sinc argument. So, recalling Eq. (3.2c.22),

$$S_{a,b}(x,y) = \text{sinc}(\tfrac{x}{h} - a)\ \text{sinc}(\tfrac{y}{h} - b). \qquad (3.2c1b.4)$$

Fig. (3.2c1b.1) shows $S_{10,10}(x,y)$ which has a maximum value of 1.0 occurring at (x,y) = (10h,10h). The object region represented is 25hx25h (25x25 grid) but has been more finely discretized to five elements per pixel (125x125 array) to indicate the sinc

function behavior more clearly. Indeed, if a 25x25 grid is used, then in some cases all pixels but one are zero because the points of evaluation all fall on zeros of the double sinc function. Yet within the numerical integration routine, sampling far finer than h/5 is used in the accumulation of the sum, so at least the h/5 sampling here roughly indicates the computational situation. Noting that the maximum value of σ in the numerical integration term is x_z/k_0, for typical sampling $h = \lambda/4$,

$$\sigma_{max} = \frac{x_z}{k_0} = \frac{\lambda x_z}{2\pi} = \frac{2x_z}{\pi}h \simeq 2.5h \, . \qquad (3.2c1b.5)$$

The extent of x_z/k_0 around the point (10h,10h) is shown in Fig. (3.2c1b.2); for all numerical integration terms this circle of integration is shifted over (and for the scattered field equations, off) the object region, and is centered on the point of evaluation of the field. Clearly, when the separation of pixel of evaluation from the pixel of summation becomes large enough, the maximum will fall outside the boundary of integration. Because the double sinc function is a factor in the integrand, for large offsets of point of evaluation with respect to point of summation, the integrand will be very small for all values of σ considered in the integrand of the finite integration term because only the greatly reduced sinc tails will fall within the integration region (again, limited to a radius of only about 2.5 pixels).

The double sinc function of Fig. (3.2c1b.1) depends upon shifts from the location of the pixel of summation in Eqs. (3.1e.22). That shift resulted from the expansion of γf over a

sinc basis on a two-dimensional sampling scheme within the object region, the region of support of γf. But the other factor of the integrand in Eq. (3.2c.17), $Y_0^E(\cdot)$, depends only upon the shifts from the location of the pixel of evaluation of the field. This latter shift, quantified by σ, is the convolutional shift in the original integral equation (Eq. (3.1c.2)). Figure (3.2c1b.3) is a plot of $-Y_{0_{10,10}}^E(x,y)$, where

$$Y_{0_{n,m}}^E(x,y) = Y_0^E[k_0 - \sqrt{(\frac{x}{h} - n)^2 + (\frac{y}{h} - m)^2}\ ]. \qquad (3.2c1b.6)$$

The lowest value of argument allowed to be sent to $Y_0^E(\cdot)$ was in this case 10^{-35}, giving a maximum value of 56 at (10h,10h) (clearly, $-Y_{0_{n,m}}^E(x,y)$ achieves its maximum value for (x,y) = (nh,mh)). In Fig. (3.2c1b.4) is shown a simultaneous display of the two previously plotted functions, which are multiplied to form the integrand. In this particular plot are shown $S_{5,5}(x,y)$ and $-Y_{0_{10,10}}^E(x,y)$; because of the separation between point of evaluation of the field and point of summation, the two maxima are separated, making the product small everywhere.

The integrand is illustrated next. Define

$$\mathcal{L}_{n,m;a,b,}(x,y) = S_{a,b}(x,y)Y_{0_{n,m}}^E(x,y). \qquad (3.2c1b.7)$$

Then in Fig. (3.2c1b.5) is plotted $\mathcal{L}_{10,10;5,5}(x,y)$. The maximum is 0.006 at (9h,9h). Note that at $\sigma = 0$, unless (n,m) = (a,b) the multiplying function $S_{a,b}(nh,mh) = 0$, so the large (in theory, infinite) value of $Y_{0_{n,m}}^E(nh,mh)$ is multiplied by zero. To see detail below the zero-value plane for a large-separation case,

Fig. (3.2c1b.6) shows $\mathcal{L}_{22,22;5,5}(x,y)$. The positive peak (maximum is 4.6x10-4 at (21h,21h)) is modulated by the sinc function fluctuations. Figure (3.2c1b.7) shows a closer approach of the point of observation to the point of summation: $\mathcal{L}_{10,10;8,8}(x,y)$. Even here the maximum value is very small in magnitude--only 0.05 at (9h,9h). But when the two points coincide, as in Fig. (3.2c1b.8) which shows $\mathcal{L}_{10,10;10,10}(x,y)$, the maximum value is in theory infinity. (Of course, if all figures here were scaled the same, this is the only figure that would show any nonzero values.)

To save computation time, only those pixels fairly close together were included in the numerical integration. It was found numerically (and can be justified considering the shape of large values of the double sinc function in Fig. (3.2c1b.1)) that a sufficient domain of evaluation was as shown in Fig. (3.2c1b.9):

$$f(\Delta x,\Delta y) = \begin{cases} 1 & |\Delta x|\leq 2 \text{ or } |\Delta y|\leq 2 \\ 0 & \text{otherwise} \end{cases}. \qquad (3.2c1b.8)$$

When larger separations were included, the modifications in the coefficients were negligible within more than three significant figures. In the final set of figures the matrix of coefficients is plotted. First is the distribution of the numerical term only (Fig. (3.2c1b.10)). As expected, it peaks for zero spatial shift. Next, in Fig. (3.2c1b.11), is the distribution of the real part of the analytically evaluated term only. These two terms sum to yield the total real part of the C_{1j}, shown in Fig. (3.2c1b.12). Note the difference from Fig. (3.2c1b.11), showing the necessity of including the numerically evaluated term. In fact, the maximum

absolute value occurring in Fig. (3.2c1b.10) is 0.018, slightly larger than that in Fig. (3.2c1b.11), -0.014, both of which occur at (0,0), the center of the plots; thus, near the origin the numerical term dominates. The imaginary part of the analytically evaluated term, which is the total imaginary part of the C_{1j}, is shown in Fig. (3.2c1b.13).

3.2c1c Exploitation of symmetry

Recalling that $1 \leq a,b,m,n \leq n_{max}$, it is clear that unless the symmetry of the integral is exploited, the computations will take extremely long even for small grid sizes, n_{max}. The above limits on a,b,m,n translate to

$$-(n_{max} - 1) \leq n - a, m - b \leq n_{max} - 1. \qquad (3.2c1c.1)$$

The symmetry in R(nh,mh,ah,bh) resides in the fact that R depends on only (m - b) and (n - a). Furthermore, the order is not important ({m - b = 7,n - a = 6} is equivalent to {m - b = 6,n - a = 7}). To compute the number of distinct integrals required, note that because of Eq. (3.2c1c.1), the following are unique combinations of {m - b,n - a} where $n' \triangleq n_{max} - 1$: (-n',-n') through (-n',+n')--2n' + 1 integrals; (-n' + 1,-n' + 1) through (-n' + 1,+n')--2n' terms; ... to term (n',n')--one term. Summing the numbers of unique coefficients gives

$$2n'+1+2n'+ \ . \ . \ . \ + 1 = (2n'+1)^2 - \sum_{n=0}^{2n'} n'$$

$$= (2n'+1)^2 - \frac{2n'(2n'+1)}{2}$$

$$= (2n'+1)(2n'+1-n')$$

$$= (2n'+1)(n'+1)$$

$$= (2n_{max}-1)n_{max}. \qquad (3.2c1c.2)$$

It follows from the above discussion that the C_{1j} coefficients should be stored in the following order, where the row number from the top corresponds to the array location for the given coefficient and where DIFF 1 and DIFF 2 are found as the minimum and maximum differences between n - a and m - b, respectively.

DIFF 1	DIFF2	
-n'	-n'	
"	"	
"	"	level $\mathcal{N}$ = 0
"	"	
-n'	-n'	
-(n'-1)	-(n'-1)	
"	"	
"	"	level $\mathcal{N}$ = 1
"	"	
-(n'-1)	n'	
"	"	
"	"	
"	"	
n'-1	n'-1	
n'-1	n'	
n'	n'.	

(3.2c1c.3)

Let $\mathcal{N}$ = DIFF 1 + n' and $\mathcal{M}$ = DIFF 2 + n'. There are 2n' + 1 levels (values of $\mathcal{N}$) above. The levels, from bottom up, have 1,2,3,... elements. Level $\mathcal{N}$ from the top is 2n' + 1 - $\mathcal{N}$ levels from the bottom. The total number of elements in these lower levels is

$$\sum_{i=1}^{2n'+1-\eta} i = \frac{1}{2}(2n_{max}-1-\eta)(2n_{max}-\eta)$$
$$= 2n_{max}^2 - n_{max} + \eta(-2n_{max} + \frac{1}{2}) + \frac{\eta^2}{2}. \qquad (3.2c1c.4)$$

The total number of distinct coefficients is found by evaluating the above for $\eta = 0$:

$$\# \text{ DISTINCT COEFF'S} = n_{max}(2n_{max}-1). \qquad (3.2c1c.5)$$

Therefore, the total number of elements above in the levels from level 0 through level $\eta - 1$ is

$$n_{max}(2n_{max}-1) - [2n_{max}^2 - n_{max} + \eta(-2n_{max} + \frac{1}{2}) + \frac{\eta^2}{2}]$$
$$= \eta(2n_{max} - \frac{1}{2}) - \frac{\eta^2}{2} \qquad (3.2c1c.6)$$

so the index for the beginning of level η is

$$\eta(2n_{max} - \frac{1}{2}) - \frac{\eta^2}{2} + 1 = \frac{\eta}{2}(4n_{max}-\eta-1) + 1. \qquad (3.2c1c.7)$$

From there, to index the correct element within that level, $m-\eta$ is added (the excess over the value at the beginning of the level). Therefore,

$$\text{INDEX} = \frac{\eta}{2}(4n_{max}-\eta-1) + 1 + m - \eta \qquad (3.2c1c.8)$$

is the index to access in the C_{lj} coefficients array. So when forming the sums in Eq. (3.1e.22), the C_{lj} will require first determining DIFF 1 = min(n - a,m - b) and DIFF 2 = max(n - a,m - b). Then realizing that, as η ranges from zero to 2n' while DIFF 1 ranges from -n' to n', η = DIFF 1 + n', and that the excess

over the beginning of the desired level is simply $\mathcal{M}-\mathcal{N}$ = DIFF 2 - DIFF 1, the correct index can now be obtained from Eq. (3.2c1c.8) above.

Thus, by exploiting symmetry of the C_{1j} the number of coefficients that must be computed has been reduced from n_{max}^4, the total possible combinations of m,p,a, and b, to only $n_{max}(2n_{max}-1)$, and the correct index is easily computable.

3.2c2 Measured scattered field equations coefficients D_{mj}.

Attention is now turned to the evaluation of the field at the receivers. The result of the first integration in Eq. (3.2c.17) is

$$-\frac{jh^2}{4}H_0^C[k_0\sqrt{(x_m-ah)^2 + (y_m-bh)^2}] \qquad (3.2c2.1)$$

where the observation point is (x_m,y_m). Because of the large argument (e.g., $(\pi/h)\cdot(x_m - ah + \sigma\cos\phi)$ where only x_m is very large) of the sinc functions in the numerically integrated term, the result is negligible and so need not be calculated. Indeed, for values of $|x_m/h - a|$ and $|y_m/h - b|$ larger than only about 2 or 3, the integration was numerically verified to be negligible within more than three significant figures. Notice that although the numerical term would indicate that the interaction of pixels with the observation point is localized (significant only if the point of evaluation is near the point of summation), the first term $((-j/4)h^2H_0^C(\cdot))$ is nonzero in all cases (whether the observation point is a grid point or a receiver) over large

distances (at 100 wavelengths, 2% of its maximum); certainly uniform in order of magnitude over the entire (if small) object region. For this reason, the sums in Eqs. (3.1e.21) and (3.1e.22) can not be truncated; the contributions from all pixels must be kept. Upon examination of Fig. (3.1e.2), it is evident that there is symmetry in the D_{mj} coefficients. For example, the coefficient for a receiver at A1 from pixel 1 should equal that for a receiver at C1 from pixel 111. Description of this symmetry is made more precise in the following discussion. (Note that an implementation not accounting for this symmetry was tried, and results were identical, but slower.)

To exploit symmetry, it is easiest if the number of receivers, n_{det}, is divisible by eight. Also, it is argued both in (Johnson and Tracy, 1983) and in Section 3.2g that n_{det} be of the same order as the number of transmitters, n_{trans}. Thus, to obtain values for n_{det} and n_{trans}, take the square root of the number of equations (determined by the number of unknowns times the overdetermination factor for the scattered field equations) and keep adding one to n_{det} until it becomes divisible by eight. Then divide n_{meas}, the total number of measurements, by n_{det} to obtain n_{trans}. Considering the geometrical symmetry in Fig. (3.2c2.1), where n_{det} has been set to 32, it is easily seen that all of the coefficients can be related to those in Section A (A1 through A5); therefore, one need calculate only those coefficients for the A region. Here the mapping from other sections to Section A will be given.

If one considers the orientation of a receiver position with respect to the square object region, one can see, for example, that B1 is related to A4 and C1 to A1. The following set of equations details this for all 32 receivers in Fig. (3.2c2.1). (32 is used here only to help show the symmetries; for an 11x11 grid, n_{det} = 16 would be used for double overdetermination, which was usually used.) Letting nd8 = $n_{det}/8$,

B1	is related to	A4	$B_i = A_{nd8-(i-1)}$
B2	"	A3	
B3	"	A2	
C1-C5	are related to	A1-A5	$C_i = A_i$
D1-D3	"	A4-A2	$D_i = A_{nd8-(i-1)}$
E1-E5	"	A1-A5	$E_i = A_i$
F1-F3	"	A4-A2	$F_i = A_{nd8-(i-1)}$
G1-G5	"	A1-A5	$G_i = A_i$
H1-H3	"	A4-A2	$H_i = A_{nd8-(i-1)}$. (3.2c2.2)

Note the vague term "is related to" used above. Now, pixel by pixel, receiver by receiver, the relation will be made explicit. A representative pixel will be analyzed. Examination of any other pixel will show that the mapping is generally valid. Let n_{max} = 11, so in the following, whenever 12 appears in a final result, it may be generalized to $n_{max} + 1$. Let i be the detector number within any given region. Pairs of numbers (l,j) refer to pixel coordinates (row,column). Also, for convenience, let i′ = nd8 - (i - 1). The letters A(·), B(·), C(·), etc., refer to the complex valued D_{mj} coefficient for region A, B, C, etc.

Section A--compute directly.

Section B. Referring to Fig. (3.2c2.1),

$B_i(m,n) \Rightarrow B_i(2,4) = A_i(8,10) \Rightarrow A_i(12 - n,12 - m)$

reverse → (4,2); subtract both from 12 (3.2c2.3)

Section C.

$C_i(m,n) \Rightarrow C_i(2,4) = A_i(4,10) \Rightarrow A_i(n,12 - m)$

reverse → (4,2); subtract 2 from 12 (3.2c2.4)

Section D.

$D_i(m,n) \Rightarrow D_i(2,4) = A_i(2,8) \Rightarrow A_i(m,12 - n)$

subtract 4 from 12 (3.2c2.5)

Section E.

$E_i(m,n) \Rightarrow E_i(2,4) = A_i(10,8) \Rightarrow A_i(12 - m,12 - n)$

subtract both from 12 (3.2c2.6)

Section F.

$F_i(m,n) \Rightarrow F_i(2,4) = A_i(4,2) \Rightarrow A_i(n,m)$

reverse (3.2c2.7)

Section G.

$G_i(m,n) \Rightarrow G_i(2,4) = A_i(8,2) \Rightarrow A_i(12 - n,m)$

reverse → (4,2); subtract 4 from 12 (3.2c2.8)

Section H.

$H_i(m,n) \Rightarrow H_i(2,4) = A_i(10,4) \Rightarrow A_i(12 - m,n).$

subtract 2 from 12 (3.2c2.9)

Now the sections are quantitatively defined: (let $1 \leq i_{det} \leq n_{det} = 32$ here only)

SECTION	i_{det}	GENERAL FORM FOR i_{det}
A	1-5	$1 \leq i_{det} \leq n_{d8}+1$
B	6-8	$nd8+2 \leq i_{det} \leq 2n_{d8}$
C	9-13	$2n_{d8}+1 \leq i_{det} \leq 3n_{d8}+1$
D	14-16	$3n_{d8}+2 \leq i_{det} \leq 4n_{d8}$
E	17-21	$4n_{d8}+1 \leq i_{det} \leq 5n_{d8}+1$
F	22-24	$5n_{d8}+2 \leq i_{det} \leq 6n_{d8}$
G	25-29	$6n_{d8}+1 \leq i_{det} \leq 7n_{d8}+1$
H	30-32	$7n_{d8}+2 \leq i_{det} \leq n_{det}.$ (3.2c2.10)

In terms of storage order of the coefficients, there are two arrays: one for the real part of D_{mj}, DR, and one for the imaginary part of D_{mj}, DI (analogously to the C_{1j}). However, here (as opposed to the C_{1j}) the arrays are two-dimensional. One subscript is related to the (row,column) double subscript in the above mapping discussion by the relation made explicit in Fig. (3.2c2.1) for each pixel. The mathematical expression for this mapping is

$$(m,n) \Leftrightarrow j = (m-1)n_{max}+n. \qquad (3.2c2.11)$$

The other subscript of the D_{mj} arrays is i or i′ of A in the above mapping discussion. Now one can explicitly equate coefficients for Sections B through H to the appropriate element of arrays DR and DI. Again noting that "12" in the coefficient mappings is generalized to n_{max} + 1,

$A_i(m,n) = DR,DI[i, (m-1)n_{max}+n]$

$B_i(m,n) = DR,DI[i', (n_{max}-n)n_{max}+n_{max}+1-m)]$

$C_i(m,n) = DR,DI[i, (n-1)n_{max}+n_{max}+1-m]$

$D_i(m,n) = DR,DI[i', (m-1)n_{max}+n_{max}+1-n]$

$E_i(m,n) = DR,DI [i, (n_{max}-m)n_{max}+n_{max}+1-n]$

$F_i(m,n) = DR,DI [i', (n-1)n_{max}+m]$

$G_i(m,n) = DR,DI[i, (n_{max}-n)n_{max}+m]$

$H_i(m,n) = DR,DI[i', (n_{max}-m)n_{max}+n].$ (3.2c2.12)

However, as the index for counting through the receivers is i_{det}, i and i' above must be translated into expressions involving i_{det}:

REGION	i	i'
A	i_{det}	-----
B	$[i_{det} - (n_{d8}+1)]$	$2n_{d8}+2 - i_{det}$
C	$i_{det} - (2n_{d8})$	-----
D	$[i_{det} - (3n_{d8}+1)]$	$4n_{d8}+2 - i_{det}$
E	$i_{det} - (4n_{d8})$	-----
F	$[i_{det} - (5n_{d8}+1)]$	$6n_{d8}+2 - i_{det}$
G	$i_{det} - (6n_{d8})$	-----
H	$[i_{det} - (7n_{d8}+1)]$	$8n_{d8}+2 - i_{det}$

(3.2c2.13)

where the expressions in parentheses are never actually used in the program, but are included for reference.

Thus, by exploiting the symmetry of the D_{mj}, the number of D_{mj} that needs to be computed has been reduced from $n_{det} \cdot N$ to $(n_{det}/8 + 1) \cdot N$ (where $N = n_{max}^2$).

3.2d Scattered field generation

Two methods are considered here:

(1) In the first method, a γ^{ex} is chosen to be reconstructed, and then the reconstruction Eqs. (3.1e.21) and (3.1e.22) are used to generate scattered field data for input to the reconstruction algorithm as described above. Using γ^{ex} in Eq. (3.1e.22) one solves for f^{ex}. Then both γ^{ex} and f^{ex} are used in Eqs. (3.1e.21) to obtain f^{sc}. The problem with this method is that, while self-consistency of the algorithm may be demonstrated, the ability to reconstruct with experimental data is not ascertained.

(2) The second method produces data which are mathematically closer to what ideally would be measured experimentally, and are obtained without using the sinc function expansion reconstruction equations. Here γ is chosen to be an object for which exact scattered data (that is, the exact solution of Eq. (3.1.1)) can be calculated. In two dimensions, the obvious choice is a cylindrically shaped object of circular cross section. The derivation for the exact scattered field, which could not be found in the literature, is derived in the following subsection.

3.2d1 Acoustic scattering of an incident cylindrical wave by an infinite circular cylinder

This subsection (published in Cavicchi and O'Brien, 1988) contains a derivation of the exact fields in fluid media associated with an infinite cylinder of circular cross section in the

presence of an incident cylindrical wave. Exact expressions for the scattered field from a cylinder exposed to a plane wave are well-known (Longley and O'Brien, 1982) and (Morse and Ingard, 1968). The problem was solved for a rigid cylinder exposed to a cylindrical wave in 1962 (Shenderov, 1962). For the electromagnetic case of a conducting cylinder, scattered fields only outside the cylinder were derived (Wait, 1952), and can be shown to be mathematically equivalent to the corresponding expression given here. Expressions for the case of acoustic scattering in elastic solid cylindrical media appear in (Fang, 1963). A result for the acoustic case (again, only for the exterior region) for fluid media appears in (Azimi and Kak, 1985), but it has an erroneous factor of j^m in the series expansion in their Eq. (2.24). (This fact is borne out analytically in the present derivation. Numerically, it was also verified, in that the scattered field expressions given in this paper computationally satisfy the condition of continuous pressure across the cylinder boundary. However, use of their Eq. (2.25), which <u>is</u> essentially correct, in their Eq. (2.24) with the analogously derived expression for the internal field including the j^m factor does not. Further numerical verification of the validity of the formulas presented here can be found in the remainder of this thesis, where use of these formulas led to successful tomographic reconstructions of cylinders.) For computer tomography simulations, it is useful to have available the scattered field at selected points outside the cylinder. Also, testing of reconstruction algorithms can make use of exact field distributions within the object region, and

consequently inside the cylinder. In fact, this solution is used for just that in Section 5.7f to explore behavior of the sinc basis moment method for large contrast cylinder reconstructions. Therefore, it is helpful to have available the total field at any desired point in space. Such expressions are derived here.

Because the circular cylinder is symmetric, clearly all that needs to be specified geometrically is the angle ϕ between transmitter and detector and the radii of the detector and transmitter measured from the cylinder center; the coordinate system can be aligned with the transmitter-cylinder center line (see Fig. (3.2d1.1)). The wave equation in polar coordinates is

$$\partial^2 f/\partial t^2 = c^2\nabla^2 f = (c^2/r)[\partial/\partial r(r\partial f/\partial r) + (1/r)\partial^2 f/\partial\phi^2]. \quad (3.2d1.1)$$

Assuming time dependence $e^{+j\omega t}$ the general solution of Eq. (3.2d1.1) is

$$f(r,\phi,t) = \sum_{m=0}^{\infty} [A_m J_m(kr) + B_m Y_m(kr)]\cos(m\phi)\ e^{j\omega t} \quad (3.2d1.2)$$

where A_m and B_m are complex weighting coefficients, $J_m(kr)$ and $Y_m(kr)$ are Bessel functions of the first and second kinds, respectively, and $\cos(m\phi)$ is the angular function. $\cos(m\phi)$ can be used instead of $e^{jm\phi}$ because the problem is symmetric with respect to the $\phi = 0$ line.

Let ρ_1 and ρ_0, c_1 and c_0, and k_1 and k_0 be, respectively, the densities, sound velocities, and wavenumbers inside and outside the cylinder. Equation (3.2d1.1) must be solved separately inside the cylinder($f = f^w$) and outside ($f = f^{sc} + f^{inc}$). In

order to match boundary conditions, the representation of the incident cylindrical wave (centered on the source) must be changed from one with the source point as the origin (in which case one has the representation $f^{inc} = H_0^{(2)}(k_0 r_s)$ to one with the center of the cylinder as the origin. Here r_s is the distance between the desired point of evaluation and the line (cylindrical point) source, and in general, $H_m^{(2)}(x) = J_m(x) - jY_m(x)$ is the mth-order Hankel function of the second kind. If ψ designates the angle between the source-cylinder center line and the source-observation point line [see Fig. (3.2d1.1)], and R the distance from the source to origin (that is, the cylinder center), the addition theorem for Bessel functions states that for any Bessel function $Z_n(kr)$,

$$e^{jn\psi} Z_n(kr) = \sum_{m=-\infty}^{\infty} J_m(kr) Z_{m+n}(kR) e^{jm\phi} \qquad (3.2d1.3)$$

where on the right-hand side r is measured in the new coordinate system. In the present case, n = 0 and Z is $H^{(2)}$ ($H_0^{(2)}$ is what is necessary to expand). Therefore, the needed specialization of Eq. (3.2d1.3) is

$$Z_0(kr) = \sum_{m=-\infty}^{\infty} J_m(kr) Z_m(kR) e^{jm\phi}. \qquad (3.2d1.4)$$

Assuming a unity amplitude source and dropping the $e^{j\omega t}$ factor, and again noting symmetry about the line $\phi = 0$, the incident field can be written as

$$f^{inc}(r,\phi) = \sum_{m=-\infty}^{\infty} J_m(k_0 r)\ H_m^{(2)}\ (k_0 R)\ \cos(m\phi). \qquad (3.2d1.5)$$

For computation, it will be convenient to express Eq. (3.2d1.5) as a sum from m = 0 to ∞, rather than from $-\infty$ to $+\infty$. Using the identity

$$Z_{-n}(x) = (-1)^n Z_n(x) \qquad (3.2d1.6)$$

which is true for n an integer, Eq. (3.2d1.5) can be rewritten as

$$f^{inc}(r,\phi) = J_0(k_0 r)\, H_0^{(2)}(k_0 R) + \sum_{m=1}^{\infty} J_m(k_0 r)\, H_m^{(2)}(k_0 R)\, [1+(-1)^{2m}]\cos(m\phi) \qquad (3.2d1.7a)$$

or

$$f^{inc}(r,\phi) = J_0(k_0 r)\, H_0^{(2)}(k_0 R) + 2\sum_{m=1}^{\infty} J_m(k_0 r)\, H_m^{(2)}(k_0 R)\cos(m\phi). \qquad (3.2d1.7b)$$

Similarly, the scattered field can be written as

$$f^{sc}(r,\phi) = \sum_{m=0}^{\infty} S_m H_m^{(2)}(k_0 r)\cos(m\phi) \quad (r \geq a) \qquad (3.2d1.8)$$

and the field within the cylinder as

$$f^{W}(r,\phi) = \sum_{m=0}^{\infty} W_m J_m(k_1 r)\cos(m\phi) \quad (0 \leq r \leq a). \qquad (3.2d1.9)$$

Now terms can be equated when the boundary conditions are enforced because all of the above expressions have the same form.

In the above, the f's denote pressure fields. Keeping this in mind in the following discussion of boundary conditions, the acoustic linear equation of motion can be written

$$\rho \partial \vec{u}/\partial t = -\nabla f \qquad (3.2d1.10)$$

where $\vec{u}$ is the particle velocity and ρ is the density. For the

two-dimensional case in cylindrical coordinates, Eq. (3.2d1.10) becomes

$$\rho \partial \vec{u}/\partial t = -(\hat{r}\partial f/\partial r + \hat{\phi}(1/r)\partial f/\partial \phi) \qquad (3.2d1.11)$$

where $\vec{u} = \hat{r}u_r + \hat{\phi}u_\phi$.

The boundary condition involving velocity is the continuity of u_r--thus, only u_r needs to be calculated and not u_ϕ. The equation for u_r is, from Eq. (3.2d1.11) and the $e^{j\omega t}$ time dependence,

$$j\omega\rho u_r = -\partial f/\partial r \quad \text{or} \quad u_r = (j/\omega\rho)\partial f/\partial r. \qquad (3.2d1.12)$$

Now the incident field radial velocity u_r^{inc} is calculated:

$$u_r^{inc}(r,\phi) = (j/\omega\rho_0)(H_0^{(2)}(k_0R)\partial/\partial r\,[J_0(k_0r)] + 2\sum_{m=1}^{\infty}\{H_m^{(2)}(k_0R)\partial/\partial r[J_m(k_0r)]\cos(m\phi)\}). \qquad (3.2d1.13)$$

Using the identity

$$d/dx[Z_m(x)] = 1/2[Z_{m-1}(x) - Z_{m+1}(x)] \qquad (3.2d1.14)$$

and Eq. (3.2d1.6),

$$\partial/\partial r[Z_0(k_0r)] = (k_0/2)\,[Z_{-1}(k_0r) - Z_1(k_0r)] = -k_0Z_1(k_0r) \qquad (3.2d1.15a)$$

$$\partial/\partial r[Z_m(k_0r)] = (k_0/2)\,[Z_{m-1}(k_0r) - Z_{m+1}(k_0r)] \qquad (3.2d1.15b)$$

so Eq. (3.2d1.13) becomes, for the incident field radial velocity,

$$u_r^{inc}(r,\phi) = -j/Z_0(H_0^{(2)}(k_0R)\ J_1(k_0r) + \sum_{m=1}^{\infty} \{H_m^{(2)}(k_0R)[J_{m+1}(k_0r) - J_{m-1}(k_0r)]\cos(m\phi)\}) \quad (3.2d1.16)$$

where henceforth Z refers to acoustic impedance so $Z_0 = \rho_0 c_0$. Use of Eq. (3.2d1.14) is made for concreteness, in that the results contain only easily obtainable functions in a computational setting.

Similarly, for the scattered field radial velocity,

$$u_r^{sc}(r,\phi) = (j/\omega\rho_0)\ \{S_0 \partial/\partial r[H_0^{(2)}(k_0r)] + \sum_{m=1}^{\infty} \{S_m \partial/\partial r[H_m^{(2)}(k_0r)]\cos(m\phi)\}\} \quad (3.2d1.17a)$$

$$u_r^{sc}(r,\phi) = -(j/Z_0)\ \{S_0 H_1^{(2)}(k_0r) + \sum_{m=1}^{\infty} \{(S_m/2)\ [H_{m+1}^{(2)}(k_0r) - H_{m-1}^{(2)}(k_0r)]\cos(m\phi)\}\}. \quad (3.2d1.17b)$$

Now the internal field radial velocity can immediately be written:

$$u_r^{W}(r,\phi) = -(j/Z_1)\ .\ \{W_0 J_1(k_1r) + \sum_{m=1}^{\infty} \{(W_m/2)\ [J_{m+1}(k_1r) - J_{m-1}(k_1r)]\cos(m\phi)\}\} \quad (3.2d1.18)$$

where Z_1 is defined as follows. In Eqs. (3.2d1.13) and (3.2d1.15), define

$$\frac{1}{Z_1} = \frac{k_1}{\omega\rho_1} = \frac{1-j\alpha_1 c_1/\omega}{\rho_1 c_1} \quad (3.2d1.19a)$$

or

$$Z_1 = \frac{\rho_1 c_1}{1-j\alpha_1 c_1/\omega} \quad (3.2d1.19b)$$

where α_1 is the pressure absorption coefficient within the cylinder.

Now the two boundary conditions may be applied:

(1) continuous pressure across the cylinder boundary r = a, where a is the cylinder radius

$$f^{W}(r=a,\phi) = f^{sc}(r=a,\phi) + f^{inc}(r=a,\phi) \qquad (3.2d1.20)$$

(2) continuous radial velocity across the cylinder boundary

$$u_r^{W}(r=a,\phi) = u_r^{sc}(r=a,\phi) + u_r^{inc}(r=a,\phi). \qquad (3.2d1.21)$$

First consider the case m = 0.

Boundary condition (1) reads

$$W_0 J_0(k_1 a) = S_0 H_0^{(2)}(k_0 a) + J_0(k_0 a) H_0^{(2)}(k_0 R) \qquad (3.2d1.22)$$

and boundary condition (2) reads (with division by -j)

$$(1/Z_1) W_0 J_1(k_1 a) = (1/Z_0)\,[S_0 H_1^{(2)}(k_0 a) + H_0^{(2)}(k_0 R) J_1(k_0 a)]. \qquad (3.2d1.23)$$

The solution for S_0 is

$$S_0 = (-1/\Delta_0)\,[J_1(k_0 a) J_0(k_1 a) - J_0(k_0 a) J_1(k_1 a) Z_r] H_0^{(2)}(k_0 R) \qquad (3.2d1.24)$$

and the solution for W_0 is

$$W_0 = (1/\Delta_0)\,[H_1^{(2)}(k_0 a) J_0(k_0 a) - H_0^{(2)}(k_0 a) J_1(k_0 a)] H_0^{(2)}(k_0 R) \qquad (3.2d1.25)$$

where

$$\Delta_0 = H_1^{(2)}(k_0 a) J_0(k_1 a) - H_0^{(2)}(k_0 a) J_1(k_1 a) Z_r \qquad (3.2d1.26)$$

where

$$Z_r = Z_0/Z_1 = \frac{\rho_0 c_0}{\rho_1 c_1}\left(1 - j\frac{\alpha_1 c_1}{\omega}\right). \qquad (3.2d1.27)$$

Note: the above formulas are identical to those for the W_0 and S_0 (m = 0 weights) for an incident plane wave (remember, the chosen incident wave is cylindrical) except for the factor

$H_0^{(2)}(k_0R)$. This is obvious because in the incident field expansion, the two zero-order terms also differ by only the same factor.

Now consider $m > 0$. The only difference between the boundary conditions for this case and the case $m = 0$ is that all zero orders are replaced by m orders and the terms from the incident field for the $m > 0$ case are twice those of the $m = 0$ case.

The solution for S_m is

$$S_m = (-2/\Delta_m)H_m^{(2)}(k_0R)\ \{J_m(k_1a)[J_{m+1}(k_0a)-J_{m-1}(k_0a)] - J_m(k_0a)[J_{m+1}(k_1a)-J_{m-1}(k_1a)]Z_r\} \quad (3.2d1.28)$$

where

$$\Delta_m = J_m(k_1a)[H_{m+1}^{(2)}(k_0a) - H_{m-1}^{(2)}(k_0a)] - H_m^{(2)}(k_0a)[J_{m+1}(k_1a) - J_{m-1}(k_1a)]Z_r. \quad (3.2d1.29)$$

Finally, the solution for W_m is

$$W_m = (2/\Delta_m)H_m^{(2)}(k_0R)\ \{J_m(k_0a)[H_{m+1}^{(2)}(k_0a) - H_{m-1}^{(2)}(k_0a)] - H_m^{(2)}(k_0a)[J_{m+1}(k_0a) - J_{m-1}(k_0a)]\}. \quad (3.2d1.30)$$

Again, the weighting coefficients differ from those for an incident plane wave by only a factor of $(j^{-m})H_m^{(2)}(k_0R)$.

3.2e Sampling density

This discussion of sampling involves only the interior object region; the transmitters/receivers are not constrained in this way. Care must be used in choosing the sampling interval h to avoid aliasing. By taking the Fourier transform of Eq. (3.1c.1) evaluated in the object region, it is seen that the

spectrum of the total field is, in two dimensions,

$$\mathcal{F}[f(\vec{r})] = \mathcal{F}[f^{inc}(\vec{r})] + \mathcal{F}[\gamma(\vec{r})f(\vec{r})] \cdot \mathcal{F}[G(r)]$$
$$= \mathcal{F}[f^{inc}(\vec{r})] + \{\mathcal{F}[\gamma(\vec{r})] ** \mathcal{F}[f(\vec{r})]\} \cdot \mathcal{F}[G(r)]. \quad (3.2e.1)$$

(The convolution in Eq. (3.2e.1) precludes use of the spatial Fourier domain to simplify calculation of high-order solutions of Eq. (3.1c.1).) In two dimensions (Mott and Massey, 1965)

$$\mathcal{F}[G(r)] = 1/(k^2-k_0^2) - j\pi\delta(k^2-k_0^2) \quad (3.2e.2)$$

which has a singularity on the k_0 circle. Its real part is shown in Fig. (3.2e.1). The incident field, being of the same form in this case as the Green function, also has the spectrum given in Eq. (3.2e.2).

The term $\mathcal{F}[f(\vec{r})]$ appears on both sides of Eq. (3.2e.1). For a rough estimation of the bandwidth of $f(\vec{r})$ it is sufficient, for weak scattering, to replace $f(\vec{r})$ by $f^{inc}(\vec{r})$ on the right-hand side of Eq. (3.2e.1). Then $(\tilde{\gamma}**\tilde{f})(\vec{k})$ will be approximately a line integral of $\tilde{\gamma}$ along a circle {of some small width due to the imaginary part of $\mathcal{F}[H_0^{(2)}(.)]$} of radius k_0 centered on $\vec{k}$. This integral will be greatest (assuming γ is a smooth bandlimited function) for $|\vec{k}|$ near k_0 ($||\vec{k}| - k_0| << \Omega_{max_\gamma}$), for then the path of integration will pass near zero spatial frequency [see Fig. (3.2e.2a)]. For large $||\vec{k}| - k_0|$, the curve of integration will pass only through high spatial frequency regions, where $\tilde{\gamma}$ is zero or negligible [see Fig. (3.2e.2b)]. The result is that $\widetilde{\gamma f}$ can be roughly approximated by a blurred spectrum of γ smeared around the k_0 circle [see Fig.

(3.2e.3)]. Depending on the magnitude of $\gamma(\vec{r})$, then, the spectrum of the total field extends out to roughly $\Omega_{max_f} = k_0 + \Omega_{max_\gamma}$ where Ω_{max_γ} is the maximum spatial frequency in γ. For γ a circular function of radius a,

$$\gamma(\vec{r}) = g(r) = \omega^2(1/c_1^2 - a/c_0^2) \cdot \begin{cases} 1 & r \le a \\ 0 & r > a \end{cases} \qquad (3.2e.3)$$

where c_1 is the speed of sound in the cylinder. The Fourier transform of $\gamma(\vec{r})$ is

$$\mathcal{F}[\gamma(\vec{r})] = \tilde{g}(\rho) = \omega^2(1/c_1^2 - 1/c_0^2)\ (2\pi a^2)J_1(2\pi a\rho)/(2\pi a\rho) \qquad (3.2e.4)$$

where here ρ is the radius of the two-dimensional spatial frequency vector (units of cylces/mm) and $J_1(\cdot)$ is the first-order Bessel function of the first kind. The envelope of $\tilde{g}(\rho)$ goes as $\rho^{-1.5}$;. At $a\rho = 1/4$, $\tilde{g}(\rho)$ is 3 dB below its maximum value ($\tilde{g}(0) = \omega^2(1/c_1^2 - 1/c_0^2)\cdot\pi a^2$). For a cylinder discretized onto a square grid, $a = m_{cyl}\cdot h$ where m_{cyl} is the number of samples in the radius of the cylinder. Then the Fourier transform of the cylinder is 3 dB below its maximum value (at $\rho = 0$) for $\rho \ge 1/(4m_{cyl}\cdot h)$, which in units of rad/mm is $\Omega_{max_\gamma} = \pi/(2m_{cyl}\cdot h)$.

The bandwidth of the Fourier transform of a product of two functions is the sum of the bandwidths of the individual functions. In the present case, the tails caused by $\gamma(\vec{r})$ in the product $\gamma(\vec{r})f(\vec{r})$ may be of substantially lower magnitude than the main component of $f(\vec{r})$, namely $f^{inc}(\vec{r})$, because of the small magnitude of $\gamma(\vec{r})$ dictated by how greatly c_1 differs from c_0. To avoid severe aliasing, one must sample at twice the maximum

frequency of γf when representing γf by a sum of its samples times shifted sinc functions. Thus, the sampling rate must satisfy

$$2\pi/h = 2\epsilon(\Omega_{max_\gamma} + \Omega_{max_f}) \quad (3.2e.5a)$$

$$= 2\epsilon(\Omega_{max_\gamma} + \Omega_{max_\gamma} + 2\pi/\lambda) \quad (3.2e.5b)$$

$$= 2\epsilon[\pi/(m_{cyl}.h) + 2\pi/\lambda] \quad (3.2e.5c)$$

or

$$h = (\lambda/2)\ (1/\epsilon - 1/m_{cyl}) \quad (3.2e.5d)$$

where $\epsilon = 1$ for 3 dB "Nyquist" sampling, $\epsilon < 1$ for undersampling, and $\epsilon > 1$ for oversampling.

Similarly, in order for the sinc function expansion of the Hankel function used in computing the C_{lj} and D_{mj} coefficients in Eqs. (3.1e.21) and (3.1e.22) to be valid, the same sampling rate, $2\pi/h$, must exceed twice the maximum frequency in that function: $\Omega_{max_{H_0}}$. The Fourier transform of the zeroth-order Hankel function, given above, falls 3 dB below its value at zero for $k = 1.55k_0$. Figure (3.2e.1) shows the real part of $\widetilde{\mathcal{F}}\{(-j/4)H_0^{(2)}(k_0r)\}$, its -3 dB frequency, and the bandwidths of γf for $m_{cyl} = 4$ and 11. However, it is actually H_0^C that is expanded, a function equal to $H_0^{(2)}$ everywhere except near the origin, where the singular Y_0 component is replaced by a smooth capping function (Tracy and Johnson, 1983) and (Section 3.2c). Thus, the actual spectrum is likely to be much narrower than that in Fig. (3.2e.1) because of the absence of the singularity, making Ω_{max_γ} the critical frequency in the determination of the sampling rate. Therefore, $\epsilon = 1$ is a rough threshold (for the

cylinder sizes considered in this study) on resulting quality of reconstructions.

If $m_{cyl} = 4$ (value for the 11x11 object region) and $\epsilon = 1$, then $h = \lambda/2.7$. For the case $m_{cyl} = 11$ (largest cylinder in this chapter), $h = \lambda/2.2$. It is stated in (Johnson and Tracy, 1983) that one must sample at $\lambda/4$. For the limit of an infinitely wide cylinder, in theory one could sample at $\lambda/2$. However, then the bandwidth of γf would approach the frequency of the singularity of the spectrum of $(-j/4)H_0^C$ so that its expansion over the sinc basis would now be invalid even though the expansion of γf may be valid. It therefore appears that the maximum sample spacing allowable is somewhere between $\lambda/2$ and $\lambda/4$.

It should be emphasized that the spectra of $\gamma(\vec{r})$, $f(\vec{r})$, and $(-j/4)H_0^C(k_0 r)$ all extend in both the x and y directions to infinite frequencies. Hence, there is no strict Nyquist frequency. As defined here, "Nyquist" frequency refers to the sampling rate above which the reconstructions are predicted to be acceptable and below which the reconstructions would be poor because of severe aliasing. The -3 dB frequency was chosen as a consistent indicator at which the slopes of the spectra are steep so that small variations in the sampling rate will show large changes in the quality of the reconstructions. Of course, in the end the simulations themselves will give the final answer. Also, for weak scattering and oversampling, the coefficients C_{1j} and D_{mj} are problem-independent because a single sampling rate will suffice for all problems likely to be encountered.

3.2f Optimization of speed of program

Two ideas that resulted in program modification to reduce run time are discussed here. Two further improvements due to parameter selection appear in Sections 3.3g and 3.3i. First, if one considers solving Eqs. (3.1e.22) by ART-type methods, one obvious procedure is to solve the subset of equations for a particular view for the internal field for a particular source position, then do the same for the next view, etc. Although in obtaining the elements of the matrix rows (in Eqs. (3.1e.22) these are $C_{lj}\gamma_j - \delta_{lj}$, δ_{lj} being the Kronekar delta function), many arithmetic operations are involved, row i is the same for each view (γ is held constant during the solution of Eqs. (3.1e.22)). Thus, it is far more efficient to update all views with row 1 (corresponding to pixel 1), then all views with row 2, etc.

Second, in ART the correction to x_j, the jth element of the unknown $\vec{x}$, is

$$x_j = x_j + \beta_1[y_i - (\vec{x},\vec{r}_i)]r_{ij}/||\vec{r}_i||^2 \qquad (3.2f.1)$$

where y_i is the ith component of the known vector $\vec{y}$ in $\vec{y} = R\vec{x}$, r_{ij} is the jth element of $\vec{r}_i$, the ith row of matrix R, and β_1 is a chosen scaling parameter. It has numerically been found that, within observable significant digits, all $||\vec{r}_i||$ are practically equal, so in the solution of Eqs. (3.1e.21) one can use $||\vec{r}_1||$ for all i. Furthermore, in the solution of Eqs. (3.1e.22) equally good results (in terms of squared error in $\tilde{\gamma}$) were obtained by using the simpler correction

$$x_i = x_i + \beta_0[y_i - (\vec{x},\vec{r}_i)] \qquad (3.2f.2)$$

where β_0 is a chosen constant. In Eq. (3.2f.2) note the simpler error scaling and the subscript i on x indicating that for row i it is sufficient to correct only the ith element of $\vec{x}$, whereas it was found that in solving Eqs. (3.1e.21) all elements of $\vec{x}$ must be corrected for each row i (see Eq. (3.2f.1)). These simplifications cut the run time in half. (It could here be argued that because the true Algebraic Reconstruction Technique corrections are not being used, none of the associated convergence properties hold. However, in an approximated sense, similar corrections are being made as in ART because row i is approximately equal to $\hat{e}_i$, where the ith element of $\hat{e}_i$ is one and the others are zero.)

Note: in the paragraph above, "equally good" was qualified with "in terms of the squared error in $\tilde{\gamma}$". Normally, a combination of $||\tilde{\gamma} - \gamma||^2$ and appearance of solution against that of the exact solution is used in assessing convergence. These are standard metrics typically used in the literature (although they are relative and therefore any thresholds will be subjectively determined). They are useful except in the gray areas between low and high accuracy; it is difficult to attempt discrete decisions about a continuum of quality.

3.2g Fractional degree of overdetermination

To reduce the deteriorating effects of measurement noise and ill-conditioning, the number of measurements can be increased by either increasing n_{trans} or n_{det}. There are N grid points, each

with a γ value and an internal field value--determined by the incident field and the scattering object. If the relevant matrix (see Eqs. (3.1e.21)) is well-conditioned (of rank N), then there are approximately N independent measurement equations because the field at every pixel makes a contribution to the scattered field at any point in space. Consequently, to obtain all the unique data and not overdetermine the system, $n_{trans} \cdot n_{det}$ would be set equal to N.

One could have one detector and N transmitters or equivalently (by reciprocity) one transmitter and N detectors. Such a setup would constitute an ill-posed problem, for then a given set of scattered field data could not be associated with a unique γ. The optimal ratio of n_{trans} to n_{det} is about one-to-one because if $x \cdot y$ is fixed, $x + y$ is minimized if $x = y$. Therefore, in this study n_{det} was chosen to be the lowest multiple of 8 greater than $\sqrt{N}$ and n_{trans} was then chosen as N/n_{det} times the overdetermination factor, plus one. Overdetermination (increasing $n_{trans} \cdot n_{det}$ above N) serves only to help average out noise present in the measured scattered field data or compensate for ill-conditioning of the matrix.

Two measures of fractional overdetermination (Q_1 and Q_2) may be defined. The first, describing the overdetermination for the entire system, compares the number of measurements to the total number of unknowns (in both Eqs. (3.1e.21) and Eqs. (3.1e.22)).

$$Q_1 = [n_{trans} \cdot n_{det} + n_{trans} \cdot N - (1+n_{trans}) \cdot N] / [(1+n_{trans}) \cdot N]$$

$$= \frac{n_{trans} \cdot n_{det}/N - 1}{1+n_{trans}} \qquad (3.2g.1)$$

where $n_{trans} \cdot n_{det}$ is the number of measurements, $n_{trans} \cdot N$ is the number of constraint equations (Eqs. (3.1e.22)) and $(1 + n_{trans}) \cdot N$ is the total number of unknowns. Note that Q_1 does not increase as n_{trans} (the number of views) gets large. For a fixed n_{det}, as n_{trans} gets large, Q_1 approaches n_{det}/N. Above, a reasonable choice for $n_{det} = \sqrt{N} = n_{max}$ was found. In any case, for a fixed n_{trans} clearly n_{det} will increase as n_{max}. Thus, as n_{trans} becomes large, Q_1 approaches a maximum of the order $1/n_{max}$; the overdetermination Q_1 declines with grid size. But then Q_1 can not also be an indicator of resulting performance, because in Section 3.3e results are shown indicating that by solving exactly the same cylindrical object reconstruction problem using different sized grids, convergence appears to be practically independent of n_{max}.

The other measure of fractional overdetermination is (consideration of Eqs. (3.1e.21) only)

$$Q_2 = n_{trans} \cdot n_{det}/N-1 \qquad (3.2g.2)$$

where N represents the N unknown γ values in Eqs. (3.1e.21). The parameter Q_2 may actually be more useful than Q_1; (very low) noise is assumed in the constraint equations, except that introduced via the estimated γ--which in turn is determined by the scattered field measurements via the solution of Eqs. (3.1e.21). The incident field, in the constraint equations, can be determined to arbitrary accuracy. Thus, the correction by overdetermination for the greatest source of noise--that present in the

measurements of the scattered field--is better represented by Q_2.

Choosing n_{det} and n_{trans} as described above, for $n_{max} = 11$ they both turned out to be 16, where for all grid sizes considered Q_2 was held nearly constant at about 1.2. For comparison, in (Tracy and Johnson, 1983), $n_{det} = n_{trans} = 17$ so that $Q_2 = 1.4$. Of course, if one already were aware of the symmetry of the circular cylinder, the number of equations could be reduced and still result in satisfactory recontructions. But for a general object, Tracy and Johnson (1983) found improvement for up to 300% overdetermination.

3.3 Computational Results

In this section, results of computer simulations using the sinc basis moment method inverse scattering algorithm (Johnson and Tracy, 1983) are presented. Further results will be given in Chapter 5; the computational results in this chapter are those of the initial study contained in the paper (Cavicchi et al., 1988). Two types of object distribution were reconstructed: a Gaussian profile infinite cylinder and a circular infinite cylinder. The purpose of the former is to relate results of the present study to those of Tracy and Johnson (1983), and the purpose of the latter is to apply the algorithm to a situation for which exact scattered field data are available for the computer simulations.

The topics to be briefly discussed here include the effects of varying several parameters of the tomography-scattering object system and a few computational issues with reference to actual program runs.

For all runs, the initial guess for both the real and imaginary parts of γ was zero, c_0 was held at 1.5 mm/μsec, the frequency was 2 MHz, and the transducer-object region center distance, R, was 10λ. Except where otherwise stated, the sampling rate scale, ϵ, was about 1.5, the level of noise added to the scattered field data was 5%, and for circular cylinder reconstructions the number of samples across the radius of the cylinder, m_{cyl}, was $2/3 \cdot (n_{max}/2 + 1)$ (so that the cylinder nearly fills up the object region--a condition which for $\epsilon = 1.5$ dictates cylinder radius a to be approximately λ for the 11x11 object region grid and 2λ for the 25x25 grid).

3.3a Scattered field generation

It was found that using Eqs. (3.1e.21) and (3.1e.22) to generate the scattered field gave essentially identical reconstructions of small, low to medium contrast objects to those obtained using the "exact" equations for the circular infinite cylinder case. This is a satisfying result, for it shows that Eqs. (3.1e.21) and (3.1e.22) are very good approximations to the corresponding exact integral equations.

3.3b Test: Gaussian object

Exactly the same two-dimensional 11x11 object distribution as that reconstructed by Tracy and Johnson (1983) was reconstructed on the author's version of their algorithm. Their computation time on a PDP11/34A computer (with floating point processor) required 7 hours with 2% noise added to the data. On the lab VAX

11/730 (1 Mbyte with floating point accelerator), for generation of scattered field plus reconstruction it took eight minutes, and only four minutes for reconstruction only (read in the scattered field) (three complete iterations), to obtain a reconstruction with slightly less error $||\tilde{\gamma} - \gamma^{ex}||^2$ than and virtually identical appearance to the one in (Tracy and Johnson, 1983) (where γ^{ex} designates the exact solution for γ). Here it was found that no significant degradation occurred for additive noise in the scattered field data up to 10%. It should be noted that the present implementation does not include the low-pass filter constraint operation referred to and applied in (Johnson and Tracy, 1983).

3.3c Transducer-object region center distance R

Varying R while maintaining the same amount of measurement information had no detectable effect other than minor roundoff errors on the reconstructed object values. Because there is no loss in this system, the choice of surface for evaluation of the scattered field is immaterial; waves propagate freely in a lossless homogeneous medium. Any convex configuration could be used for detection. This fact was numerically verified; reconstruction quality was independent of R for R = 10λ, 100λ, even 1000λ.

3.3d Frequency

As noted in (Cavicchi and O'Brien, 1985), there is no effect on field distributions (except possibly scaling) if the frequency is changed but the object size in wavelengths is held constant. In fact, one can see that the coefficients S_m, etc. in Section

3.2d1 depend only on ka where k is the wavenumber in the cylinder (and kR, but this issue was just dealt with above) so that, other than scaling, the form of the fields is the same for two frequencies if ka is held constant. For this algorithm, the above fact caused object reconstruction values to be essentially identical (within roundoff errors) for the cases frequency = 2 MHz and 5 MHz, while the sampling rate scale, ϵ, was held at 1.5.

3.3e Object contrast

In Fig. (3.3e.1) an example is shown, for a 25x25 grid, of a three-dimensional perspective view of the reconstructions of speed of sound and absorption distributions for a 5% contrast (speed of sound inside the cylinder, c_1, was 5% above c_0), k_0a = 13.6 circular cylinder after four iterations. For the remaining reconstructions in this section, only center-line profiles are included, as they are more quantitative than are the perspective views because of the possibility of using labeled axes. Also, for Fig. (3.3e.1b) and all subsequent plots in this chapter, straight lines have been drawn between points (representing either pixel values or results of entire reconstructions); points have been left out for clarity.

Good reconstructions were obtained for cylindrical objects in which the speed of sound, c_1, was different from c_0 by less than about 5%. Figure (3.3e.2) is a composite showing reconstructions (heavy line) against the exact solution (lighter line) for a range of -10% to 10% for three sizes of object radius. These sizes were defined by choosing $m_{cyl} = 2/3 \cdot (n_{max}/2 + 1)$ for n_{max} =

11, 17, and 25. Looking within one column pair, one sees the dependence of reconstruction quality of speed of sound and absorption distributions upon percent object contrast for a given object radius. Within one plot, the results of both the first and fourth iterations are included. It is evident that the fourth iteration is significantly better than the first, which is the Born approximation because the internal field is set equal to the incident field. For the first iteration the reconstructed speed of sound within the cylinder is typically too close to c_0, while in the fourth iteration it oscillates about the exact solution. This "lost energy" in the real part of γ is mixed into the imaginary part, as evidenced by the erroneous cylinder shape in the reconstructed absorption distribution (which should be zero everywhere). This effect is also evident in Soumekh and Kaveh's (1986) reconstructions obtained using the Born approximation but a different method of solution. Again, in the fourth iteration the reconstructed absorption distribution oscillates about the exact solution, zero.

Beyond ±5% contrast in this implementation, convergence quality exponentially deteriorates (see Fig. (3.3e.3)), although in Fig. (3.3e.2) the error in the fourth iteration exceeded that in the first for only one case: reconstruction of absorption; -10% contrast, 25x25 grid. In all other cases, the line closer to the exact solution is iteration 4. All of the results in Fig. (3.3e.2) are plotted on a single graph in Fig. (3.3e.3), where here the abscissa is γ (not percent contrast). The ordinate is the log of the normalized error, defined as follows. The number

of pixels for which $\gamma \neq 0$ is called $n_{\gamma\neq 0}$. (Although in practice $n_{\gamma\neq 0}$ would not be known beforehand, it could be approximated by N; but because it is known here and helps to accurately characterize $\tilde{\gamma}(\vec{r})$, it is used in this chapter.) The normalized error is then the sum of the squares of $\tilde{\gamma}_j - \gamma_j^{ex}$, divided by $n_{\gamma\neq 0}$.

Table (3.3e.1) shows the results of varying object contrast of the cylinder; it contains data about all the reconstructions in Fig. (3.3e.2) and some additional runs. In Table (3.3e.1), the characteristics given for each run include percent contrast 100 $(c_1-c_0)/c_0$, c_1, the squared norm of γ divided by $n_{\gamma\neq 0}$, iter0,1,4 = the sum of squares of the error in $\tilde{\gamma}$ for iterations 0 (iteration 0 is before any correction, so $\tilde{\gamma}$ is the initial guess: zero), 1, and 4, the sum of squares of the error in iteration 4 divided by $n_{\gamma\neq 0}$, the improvement factor over the Born approximation, that is, iter 0/iter 4, and the percent error in the average value of the reconstructed c_1 within the cylinder 100 $(\tilde{\tilde{c}}_1^4 - c_1)/c_1$. A few examples of equal norm of γ for positive and negative contrasts are included; for example, -13.3% can be compared with +22%, -10% with +14.3%, etc. Each of these runs corresponds to a single point on one of the curves in Fig. (3.3e.3). A striking feature of this table is the extremely low percent error in the average speed of sound in the cylinder: less than one percent error for object contrast magnitudes of less than 10% for all object sizes considered, and less than 0.1 percent for contrast magnitudes of less than or equal to 5%. One conclusion to be drawn here is that the maximum speed of sound in the cylinder for which successful reconstruction is possible

declines with the object size, which varies with m_{cyl}. This relation will be made specific in Chapter 5 in the discussion of phase shifts through the scattering object. It should be mentioned, however, that in soft tissue the percent contrasts are entirely within these limits for these small object sizes, and the sharp edge of the cylinder function reconstructed here mimics well the boundaries found in tissue, such as blood vessels. Thus, it may be possible, using larger grid sizes, to successfully reconstruct tissue objects of more practical sizes than those considered here. In a later study (see Chapter 5), this was tried with the aid of a minisupercomputer, and the results there indicate a more pessimistic conclusion: a nonuniqueness problem requiring unreasonably good initial guesses for the field and for γ.

The dependence of convergence on n_{max} for reconstruction of identical objects was earlier denied (in Section 3.2g). Table (3.3e.2) demonstrates this for three values of object contrast within otherwise identical cylinders. The same reconstruction performed on both the 11x11 and 25x25 grids yielded equivalent results, which are characterized similarly to those in Table (3.3e.1). Thus, the fact that runs obtained using different sized grids and presented in Figs. (3.3e.2) and (3.3e.3) and Table (3.3e.1) show a decline of maximum contrast for successful reconstruction with cylinder radius is apparently not due to the use of different sized grids for the different radii.

The limitation concerning object contrast may be due to several factors, though in Chapter 5 a fundamental and overriding

factor will be proposed. In this section, three conjectured reasons for divergence occurring for high contrast objects will be discussed, as understood at the time of the writing of (Cavicchi et al., 1988). The first and third of these reasons are contained within but will be expressed in terms of the argument given in Chapter 5 for the fundamental problem-- the phase shift through the object. Thus, the points made in this section are helpful for an overall appreciation of competing limitations of the sinc basis algorithm but are subordinate to the position taken in Chapter 5. The first conjectured reason for failure in reconstructing high contrast objects is the somewhat inefficient procedure of holding one of the two multiplied unknowns constant while iterating only upon the other and vice versa; if the initial guess is too far off then there may be difficulty in getting "on track." This phenomenon may be similar in effect to the difficulty encountered in the calculation of higher-order terms of the Born Series (Slaney and Kak, 1985). This difficulty may be circumvented by further developments in the algorithm as, for example, indicated in (Johnson and Tracy, 1983).

The second factor is the aliasing of the larger moduli spatial frequencies of the sampled product γf in low speed of sound, high contrast objects. The restriction in the statement of this factor to <u>low</u> speed of sound (lower than that of the coupling medium) objects arises because: if $c_1 > c_0$ the wavelength in the object is <u>larger</u> than λ_0. In such a case, the sampling requirement in a hypothetical homogeneous medium with speed of sound c_1 would be less stringent than the case $c_1 < c_0$, where the sampling

requirement would be more stringent. However, in this study the sampling density was held far greater than what these considerations would dictate, so it is unlikely that this factor is the cause of the difficulty in reconstructing high contrast objects.

Convergence is obtainable for greater positive percent contrasts than negative percent contrasts. However, it must also be recognized that the value of γ is asymmetric with respect to c_1/c_0: $\gamma_{cyl} = \omega^2(1/c_1^2 - 1/c_02)$ so that γ^2 is larger for a given negative percent contrast than the corresponding positive percent contrast (see Fig. (3.3e.4)). This consideration leads to the third factor that may be the root of the problem: the norm (or sum of the squares of the samples) of γ. The maximum speed of sound in the Gaussian object in (Tracy and Johnson, 1983) can be raised higher than the maximum in a cylindrical object with $m_{cyl} = 2/3\cdot(n_{max}/2 + 1)$ and still obtain convergence, because $\Sigma\gamma_{gauss}$ will be less than $\Sigma\gamma_{cyl}$ due to the Gaussian dropoff. In fact, for the cylinder having (uniform) speed of sound equal to that at the center of Tracy and Johnson's (1983) Gaussian, the reconstruction error after four iterations was over two orders of magnitude higher than that of the Gaussian. This is a greater difference than one might estimate on the basis of the Gaussian being a smoother function (narrower bandwidth) than the cylinder. The advantage of considering $||\gamma||$ rather than c_1 in discussions of convergence is also evident in Fig. (3.3e.3), which exhibits a much higher degree of symmetry about the abscissa (γ) than does a similar graph with percent contrast on the abscissa. Thus, a likely cause of convergence difficulty is that when using ART to

solve large, nonsparse matrices, convergence degrades with the norm of the unknown.

But an overruling cause for difficulty is the phase shift through the object compared with that through the homogeneous coupling medium. For details of this argument, see Chapter 5.

3.3f Object size

Again, in this section the reasons for observed behavior remain as originally presented in (Cavicchi et al., 1988) and indicate factors influencing and limiting the quality of reconstructions using the sinc basis moment method. However, the severe restriction on the maximum grid size (25x25) in that study did not reveal the reason for failure that neatly encompasses effects of both object contrast and size-- the object phase shift described in Chapter 5. In the previous subsection, $||\gamma||$ increased for increasing speed of sound contrast, while within one column pair of Fig. (3.3e.2) $n_{\gamma\neq 0}$ remained constant. Actually, in the study of reconstruction of circular cylinders, the parameter best correlating (inversely) with quality of reconstructions seems to be $||\gamma||^2/n_{\gamma\neq 0}$. Changing object size for a given contrast increases both $||\gamma||$ and $n_{\gamma\neq 0}$ so that $||\gamma||^2/n_{\gamma\neq 0}$ remains constant. It was found that when the object contrast was held constant and the cylinder radius was increased from 0.9λ to 3.0λ, the reconstruction quality remained about the same (see Fig. (3.3f.1)). Note in the $k_0a = 18.9$ object that the reconstruction is somewhat smoother than those of the smaller

objects. At least two competing trends are present when increasing the cylinder radius: the discretized cylinder shape approaches more closely the smooth circle of a continuous cylinder while the sampling (for constant ϵ) approaches the singularity of $\{(-j/4)H_0^C\}$ at k_0.

Table (3.3f.1) summarizes the reconstruction results for the study of varying the object radius as shown in Fig. (3.3f.1). Here the first three columns are $n_{\gamma\neq 0}$, m_{cyl}, and k_0a; otherwise the same result characteristics are given here as in Tables (3.3e.1) and (3.3e.2). The above statement that the quality of reconstruction is relatively constant for the investigated sizes of cylinders is evident by examining the improvement factor defined earlier which, however, is beginning to decline for the largest cylinder, $k_0a = 18.9$. Also, $||\gamma||^2/n_{\gamma\neq 0}$ is essentially constant for all sizes considered, as indicated above, while in Tables (3.3e.1) and (3.3e.2) that ratio varied inversely with reconstruction quality, in terms of the squared error in $\tilde{\gamma}$.

One can also view in Fig. (3.3e.3) how increasing object size by a factor of 2.3 affects the dependence of convergence on $||\gamma||$. While for low contrast cylinders the error after four iterations is lowest for the largest cylinder, the opposite is true when $|\gamma| > 10(\text{rad/mm})^2$ (that is, object contrast greater than about 8% and less than -7%). Thus, while for larger objects the maximum contrast allowable for successful reconstruction declines, the performance for low contrast objects is actually slightly better.

3.3g Sampling density

The numerical consequences of the earlier discussion on sampling are evident in Fig. (3.3g.1). Holding everything else constant but the sample spacing h (and consequently the cylinder radius is not strictly constant because m_{cyl} was also held constant within a single plot), convergence for a -5% speed of sound mismatch is possible for ϵ above 1.0 but not for ϵ below 1.0. Note that for ϵ = 0.8-0.9, the aliasing is so great that the cylinder shape is almost lost in the reconstruction, while for ϵ greater than 1.0, the reconstructions all follow the cylinder shape and values fairly well. This being true for three different object sizes and grid sizes, as demonstrated in Fig. (3.3g.1), gives substantial credibility to the "Nyquist rate" defined in Section 3.2e. It should be stressed that the four problems being solved within one plot are not exactly the same; that is, as h varies (m_{cyl} constant) the cylinder radius varies. This was unavoidable for the 11x11 grid, for the following reason. To maintain a constant radius in the 11x11 case, for example, would, for the range of sampling rates being investigated, require m_{cyl} to be fractional (it must be integral), larger than n_{max} (in which case the cylinder would extend outside the object region), or be ≤ 2. But the important point is that convergence is being examined for (similar) situations for which h = $\lambda/2$, $\lambda/2.5$, $\lambda/4$, and $\lambda/6.5$. Table (3.3g.1) summarizes this sampling rate study. Similar information is provided to that in Tables (3.3e.1), (3.3e.2), and (3.3f.1), except that here the first columns are λ/h, ϵ, and k_0a. Of special interest here is the huge increase in

improvement factor between the runs with ϵ less than 1 and greater than 1, for all grid sizes (and radii).

Holding the radius constant at 0.9λ was tried for the 25x25 case. The experimental result for increasing the sampling rate from 1.4 to 2.1 times the "Nyquist rate" (and therefore m_{cyl} increased from 4 to 6) was that while there is an order of magnitude of improvement in the average value of the reconstructed speed of sound within the cylinder, c_1, (from -0.05% error to -0.005% error) the sum of squares of $\tilde{\gamma}_j - \gamma_j^{ex}$ relative to the number of pixels within the cylinder decreased only slightly. Of course, the improvement in $\tilde{c}_1$ may be in part due to the increased number of averaged points within the cylinder. The fact that continuing to increase the sampling rate produces diminishing returns or even degradation (examine the column "(iter 4)/$n_{\gamma\neq 0}$" in Table (3.3g.1)) may stem from the various reconstruction equations being "too close" to each other, resulting in numerical problems ((Kogan and Lopes, 1985) and (Ekstrom, 1973)). Further experimentation concerning this topic really requires larger n_{max} (grid sizes) but at the time of this study the algorithm was too slow to go much beyond 25x25 on the lab's computer.

As an aside, one might ask what the required sampling rate would be if, instead of using the -3 dB point, the zero-crossing point of $J_1(2\pi a\rho/(2\pi a\rho)$ (Jinc$(2\pi a\rho)$) had been used. A plot of Jinc$(2\pi a\rho)$ vs. $a\rho$ is shown in Fig. (3.3g.2). From the graph, the first zero-crossing occurs for $a\rho \simeq 0.61$. So $a = m_{cyl} \cdot h$ gives

$$\rho max_{\gamma} = \frac{0.61}{m_{cyl} \cdot h} \qquad (3.3g.1)$$

or

$$\Omega max_{\gamma} = 2\pi \rho max_{\gamma} = \frac{1.22\pi}{m_{cyl} \cdot h} = \frac{\pi}{2(0.41 m_{cyl})h} \qquad (3.3g.2)$$

so that in Eq. (3.2e.5d) one merely replaces m_{cyl} by $0.41 m_{cyl}$:

$$h = \frac{\lambda}{2} \left(\frac{1}{\epsilon} - \frac{1}{0.41 m_{cyl}}\right). \qquad (3.3g.3)$$

For the case $m_{cyl} = 4$, $\epsilon = 1$, h by the above equation is calculated to be $\lambda/5.1$ and for $m_{cyl} = 11$, $\epsilon = 1$, $h = \lambda/2.6$. These values of h do not correlate well with threshold behavior of reconstruction quality (the sampling threshold as defined in Section 3.2e), as did use of the -3 dB point, so the -3 dB criterion was maintained.

3.3h Iteration duration

In Herman (1975), it is stated that about 3 or 4 runs through the equations (e.g., Eqs. (3.1e.21) and (3.1e.22)) for the case of complete x-ray projections were needed for convergence. This idea was tried in the present implementation. Indeed, the error $||f^{sc} - \Sigma D\ \tilde{\gamma}\ f||^2$ in Eq. (3.1e.21) decreased for a particular complete iteration if more runs through the equations were allowed. However, two points should be raised. First, the scattered field estimation error for any given iteration is not the error most important to minimize. It is desired to minimize $||\tilde{\gamma} - \gamma^{ex}||^2$ over a <u>series</u> of iterations. Hence, it is a waste