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).

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

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

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):

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

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):

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

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

of time to try to make a particular estimate of f^{SC} "too good" at any given point in the reconstruction process. The other fact is that, in ART, an example of a "row-action" method, only one row of the relevant matrix is held in memory at a time. But if further (more than one) runs through the equations are desired per iteration, the matrix elements must be recomputed (or else reread in), which takes large amounts of time and only achieves small refinements on the estimate of only one of the two unknowns. Four runs through Eqs. (3.1e.21) and five runs through Eqs. (3.1e.22) per iteration were used over four complete iterations. The results were essentially the same as for the situation of only one run through each of Eqs. (3.1e.21) and (3.1e.22) per iteration used over four iterations. This is true even though within each iteration, $||f^{SC} - \Sigma D\ \tilde{\gamma}\ f||^2$ was substantially smaller in the former case than in the latter. In fact, as a result of this, it was realized that further time could be saved by dispensing with the calculation of the above error, which took much time to compute and is not an indicator of reconstruction quality anyway.

3.3i Relaxation parameters

The parameters β_0 and β_1 in Eqs. (3.2f.1) and (3.2f.2), respectively, scale the correction made on the relevant unknown. They were held constant because the number of iterations was only four. In iterative numerical studies, such parameters must be determined by a search method; that is, trial and error. As noted in Section 3.1i, the identification of phenomena requiring

low relaxation constants is complicated and needs further study. Suffice it here to say that in the presence of both inexactly matching discretized equations and the imperfect initial estimate of the field, underrelaxation allows convergence over a series of "alternating variables" iterations to a desired local solution. (See also the discussion of relaxation constants in Section 3.1i.)

An example of this type of search is illustrated in Fig. (3.3i.1), where the error $(1/n_{\gamma\neq 0})\cdot||\tilde{\gamma} - \gamma^{ex}||^2$ vs. β_1 is plotted for several trial runs. In this case the optimum β_1 is 0.23; however, the curve is somewhat problem-dependent. Figure (3.3i.1) was obtained for a -10% contrast object with $n_{max} = 11$. For a 10% contrast object the optimal value of β_1 shifts to about 0.29, so an average between the two would be used in practice.

A more detailed study of dependence on β_1 is shown in Fig. (3.3i.2). Here the squared error in $\tilde{\gamma}$ was plotted for each row of correction in the first four iterations (a total of 5088 points) for four values of β_1. Values of β_1 both above and below the value 0.2 gave higher errors both in the first iteration and in later iterations, though the dependence was not too sensitive over a factor of five increase over the lowest value tried. This plot is further numerical evidence for choosing a low relaxation constant for the scattered field equations.

3.3j Execution time

Taking these programming considerations into account, the tomography programs were run for four iterations. A marked

improvement is apparent from iteration 1 to iteration 4. In this study, continuing the iterations beyond four had the effect of either minor improvement on the converging reconstructions or further deterioration in the diverging reconstructions. The execution times for 11x11, 17x17, and 25x25 grid sizes were, respectively, 5 minutes, 43 minutes, and 4.5 hours on the lab's computer (VAX 11/730). The relation between these times follows almost exactly n_{max}^5, as predicted in (Johnson and Tracy, 1983). Though there are ways in which execution times could be reduced using further program optimization (see Chapter 4, for example), the computational complexity of this algorithm is one of its most serious limitations at present.

3.4 Summary

The moment method presented by Johnson and Tracy (1983) and demonstrated here and in (Tracy and Johnson, 1983) has been shown to be a successful way to solve small inverse scattering problems of the type encountered in ultrasonic computed tomography. Not only is the method self-consistent (can go from γ^{ex} to scattered fields and back to $\tilde{\gamma} \simeq \gamma^{ex}$) but it reconstructs circular cylinders equally well when given <u>exact</u> scattered field data. The main limitations of the algorithm are execution time (and storage requirements) and object contrast. The former problem will be partially solved in Chapter 4 and the latter problem will be addressed in Chapter 5. Execution time, originally thought to be the most restricting limitation on the sinc basis method, is greatly reduced (though still quite restricting) by use of the

FFT in the internal field equations convolutions, and prospects on that front look brighter with the ever-improving computational facilities increasingly available. But the mathematical/physics problem will be far more difficult to surmount, because while by clever means the time required for a set of computations can be greatly reduced, no amount of ingenuity can coerce nature into becoming describable in terms more convenient for us in our striving towards a desired end point.

CHAPTER 4

USE OF THE FFT IN THE INTERNAL FIELD EQUATIONS

4.0 Introduction

Emphasis in Chapter 3 was upon exposition of the theory and demonstration of behavior of the sinc basis moment method for small scattering cylinders. Also, several aspects of the computer program for that algorithm were optimized, and as it stands there is little room for obvious further reduction of the computations, except for the use of the fast Fourier transform to be described in this chapter. First, a discussion of the order of computation of the moment method is given, followed by arguments against reduction of that order by using the FFT for the scattered field equations convolutions. Next, the form of the discretized sums of all the field equations is shown to be convolutional. In Section 4.3 the method of using the FFT to perform convolutions is described both mathematically and pictorially. Finally, in Section 4.4, details of numerical implementation are given, and the method is shown computationally to be equivalent to performing the convolutions in the spatial domain.

4.1 Order of Computation

As noted in Chapter 3, one of the most severe limitations at present of the sinc basis moment method is its n^5 order of operations dependence, which is higher than that of any of the other algorithms described in Chapter 2. This n^5 dependence unfortunately means that a 100x100 reconstruction would take 100,000

times as long to run as does a 10x10. (Given that on the VAX 11/730 an 11x11 reconstruction takes about five minutes to run, a 100x100 would take nearly a full year!) Further, this limits the size of an object that can be reconstructed on the VAX to about a few milimeters. Although faster computers with larger storage capabilities will in the future improve the situation, the n^5 dependence must be brought down for it to be possible to bring the object region up to a practical size.

The n^5 dependence occurs for the internal field constraint equations; for the scattered field equations the order of operations dependence is n^4. The convolutions in the internal field equations may be performed more efficiently by way of the FFT. Specifically, those convolutions are two-dimensional n^2 by n^2 discrete convolutions (see Section 4.2), with a total number of operations of the order n^4. With the FFT the number is brought down to the order $2n^2\log(n^2) \sim n^2\log(n)$. The additional factor n that brings the original order to n^5 is due to the necessity of having a number of views (transmitters) of the order of n and is present no matter how the convolutions are performed.

It is intimated in (Johnson et al., 1984), that FFTs could also be used to perform the convolution in the scattered field equations. (At the time of writing there are no numerical studies anywhere that use the FFT to solve either set of equations.) Use of the FFT for these convolutions would bring the order of the scattered field equations down again from n^4 to $n^2\log(n)$. They acknowledge that to use the FFT in the scattered field equations,

the scattered field must somehow be available on the same rectangular grid as that of the object region. Yet there seemed to be confusion over even this (Johnson, private communication). That is, Johnson suggested interpolating the scattered field values onto a rectangular grid containing both the object region and the transducers. But this is obviously not at all practical, for the scattered field is known only on a ring of discrete points and the number of pixels on a side of such a discretized region, with pixels all of spacing h, would be enormous.

Subsequently, the discussion centered upon an idea raised in their paper (Johnson et al., 1984) that, according to the integral formula of Helmholtz and Kirchoff, the field at any point can be evaluated from information known on a surface enclosing the points of evaluation. That is,

$$f^{sc}(\vec{x}) = \iint_S \{f^{sc}(\vec{z})\left[\frac{\partial}{\partial n}G(|\vec{w}|)\right]\Big|_{\vec{w}=\vec{x}-\vec{z}} - G(|\vec{x}-\vec{z}|)\frac{\partial}{\partial n} f^{sc}(\vec{z})\}dS_z \tag{4.1.1}$$

where $\vec{x}$ is a grid point, $\vec{z}$ is a point on the surface of evaluation, and G is the free-space Green function. Unfortunately, $\partial f^{sc}/\partial n$ is not known. Also, it is doubtful that the relation will be unique for limited viewing angles (surfaces that do not fully enclose the object region) and furthermore, the field is not known on a continuous surface (curve, in two dimensions) but only at a small number of discrete points. The problem of not knowing the normal derivative of the scattered field is supposedly surmounted by their following statement: "In our case, the nearest grid point $\vec{x}$ to the detector points $\vec{z}$ will be used

so that $|\vec{z} - \vec{x}|$ is always small; thus, in the vicinity of $\vec{x}$, the detector looks like a plane." And so the image-pair Green function for the case of incident radiation on a finite aperture in an infinite, opaque plane screen is used in Eq. (4.1.1) because it is zero on the surface so that the normal derivative of the scattered field need not be known. But $|\vec{z} - \vec{x}|$ <u>is</u> large compared with a wavelength for any practical tomographic configuration and a detector certainly does not look like a plane. Indeed, in this study it is a point (in two dimensions--that is, a line detector). Regardless, the problem of knowing the scattered field (equivalent to knowing the total field because f^{inc} is known) is almost as hard as the original problem! If the field in the object region were known, then only the scattered field equations would have to be solved for γ, a relatively easy problem--no harder than using the Born approximation. Because the scattered field is unknown in the region where the object function is nonzero, it appears that the FFT can not be used in the scattered field equations. In a subsequent private communication, Berggren claimed that use of the FFT in the scattered field equations was now operational in his studies, provided that the scattered field is known at all points on the perimeter of the object region. It may well be that some of the objections raised above could be reduced or eliminated were this a physical possibility, but it would appear unlikely that such measurements would be accessible and obtainable in any practical setting; only for the case of computer simulation are field points available at any desired locations and spatial densities. And if, as in a yet

later conversation, this has been extended to a quasi-propagation out to the normal receiver locations, some questions still remain. For example, how do the discrete nature of the representation and the limited angle availability degrade the scattered field estimations in comparison with the error in the estimations of γ and f^{int}, and the applicability of the theorems upon which this method depends (derived for continuous functions).

Also, a serious independent attempt was made to formulate the scattered field equations in a Fourier-domain fashion with an eye to using the FFT. Problems were (1) to obtain the scattered field would require R/d (d = object region dimension) times as high resolution as at present (i.e., the function being represented in the Fourier domain would extend out to k_0R/d instead of out to k_0) and (2) one would obtain by such calculations the scattered field on a huge rectangular grid, even though it is desired only on a curve in two dimensions.

4.2 The Convolutional Form of the Matrix Coefficients

To use the FFT to perform the convolutions in the internal field equations, one must understand the behavior of the discrete coefficients C unshifted, as opposed to their shifted versions appearing in the internal equations, marked by the double subscript lj. To be general, for the present both D_{mj} and C_{lj} are considered because the discussion will explain a previously used feature of the D_{mj}-- of the two additive integral terms into which the coefficients are decomposed, only the term that can be evaluated in closed form is significant. It will also be shown

that both the C_{1j} and D_{mj} are in fact convolutional, which may not be immediately obvious. The coefficients consist of two integral terms, as presented in (Tracy and Johnson, 1983):

$$W_j(x,y) = \frac{j}{4} \int_{-\infty}^{\infty}\int H_0^C(k_0\sigma)\,\mathrm{sinc}\left(\frac{x'}{h} - a\right)\mathrm{sinc}\left(\frac{y'}{h} - b\right)dx'dy'$$

$$- \frac{jj}{4} \int_{-\infty}^{\infty}\int Y_0^E(k_0\sigma)\,\mathrm{sinc}\left(\frac{x'}{h} - a\right)\mathrm{sinc}\left(\frac{y'}{h} - b\right)dx'dy'$$

$$\text{where } r = \sqrt{(x-x')^2 + (y-y')^2}. \qquad (4.2.1)$$

For both C_{1j} and D_{mj} the first term H_0^C is expanded over a sinc basis (see Section 3.1e):

$$H_0^C\left(k_0\sqrt{(x-x')^2 + (y-y')^2}\right) \simeq$$
$$\sum_c \sum_d H_0^C\left(k_0\sqrt{(x-ch)^2 + (y-dh)^2}\right)\mathrm{sinc}\left(\frac{x'}{h} - c\right)\mathrm{sinc}\left(\frac{y'}{h} - d\right). \qquad (4.2.2)$$

In this expansion, use is tacitly made of the fact that in the first term integrals above, (Eq. (4.2.1)), the sinc functions quickly drop off outside the object region so that including only c and d that are less than n will suffice for an adequate representation of H_0^C. (If the arguments

$$\sqrt{(x-x')^2 + (y-y')^2} \quad \text{and} \quad \sqrt{(x-ch)^2 + (y-dh)^2} \qquad (4.2.3)$$

of the H_0^C in Eq. (4.2.2) are troublesome or unfamiliar in this context of a two-dimensional sinc function expansion, the H_0^C can be thought of in terms of an equivalent function defined in the usual rectangular fashion, but of course the results would be identical, though less conveniently expressed.)

The important result of this expansion is the resulting use of the orthogonality of sinc functions (see Section 2.3a) to show that the first term is

$$\frac{j}{4}h^2 \, H_0{}^C(k_0\sqrt{(x-ah)^2 + (y-bh)^2}) \tag{4.2.4}$$

which depends only on x - ah and y - bh; that is, this term is convolutional. The first integral terma can be generalized as

$$W^1(x,y) = \frac{j}{4}h^2 \, H_0{}^C(k_0\sqrt{x^2 + y^2}) \tag{4.2.5a}$$

and specialized to the internal field equations as

$$C^1_{lj} = \frac{j}{4}h^2 \, H_0{}^C(k_0h\sqrt{(n-a)^2 + (m-b)^2}) \tag{4.2.5b}$$

where the coordinates of pixel l are (nh,mh) and those of pixel j are (ah,bh). The first integral term specializes to the scattered field equations as

$$D_{mj} \simeq D^1_{mj} = \frac{j}{4}h^2 \, H_0{}^C(k_0h\sqrt{(\frac{x_m}{h} - a)^2 + (\frac{y_m}{h} - b)^2}) \tag{4.2.5c}$$

where the coordinates of transducer m are (x_m, y_m).

The second term (which will be seen to be significant only for C_{lj}) can be written generally in polar form as

$$W^2(x,y) = \frac{1}{4}\int_0^{2\pi}\int_0^{\left(\frac{x}{k_0}\right)} Y_0{}^E(k_0\sigma)\,\mathrm{sinc}(\tfrac{x+\sigma\cos\phi}{h})\,\mathrm{sinc}(\tfrac{y+\sigma\sin\phi}{h})\,\sigma d\sigma d\phi \tag{4.2.6}$$

and, specialized for the internal field equations, is

$$C_{1j}^2 = \frac{1}{4}\int_0^{2\pi}\int_0^{\left(\frac{x_z}{k_0}\right)} Y_0^E(k_0\sigma)\,\mathrm{sinc}(n-a+\tfrac{\sigma}{h}\cos\phi)\,\mathrm{sinc}(m-b+\tfrac{\sigma}{h}\sin\phi)\,\sigma\,d\sigma\,d\phi \qquad (4.2.7)$$

which can be evaluated numerically.

In the second integral term for the scattered field equations, though σ is small, x and y are very large because they represent the distances in the x- and y-directions from the detector to a point in the object region. Hence, both multiplicative sinc terms are essentially zero. Thus, the specialization of W^2 (Eq. (4.2.6)) to the scattered field equations is clearly negligible ($Y_0^E = 0$ for $\sigma > x_z/k_0$ is the reason for the upper limit on σ).

For the internal field equations, although

$$\sigma = \sqrt{(x'-nh)^2 + (y'-mh)^2} \qquad (4.2.8)$$

includes (nh,mh) and <u>not</u> (ah,bh), it becomes, in the integral

$$\frac{1}{4}\int_{-\infty}^{\infty}\int Y_0^E(k_0\sigma)\,\mathrm{sinc}(\tfrac{x'}{h} - a)\,\mathrm{sinc}(\tfrac{y'}{h} - b)\,dx'dy' \qquad (4.2.9a)$$

merely a dummy integration variable, so that when changed to polar form,

$$\frac{1}{4}\int_0^{2\pi}\int_0^{\left(\frac{x_z}{k_0}\right)} Y_0^E(k_0\sigma)\,\mathrm{sinc}(n-a+\tfrac{\sigma}{h}\cos\phi)\,\mathrm{sinc}(m-b+\tfrac{\sigma}{h}\sin\phi)\,\sigma\,d\sigma\,d\phi \qquad (4.2.9b)$$

it is only the relative distances in the x- and y-directions between point of evaluation (pixel l) and pixel of the summation performed over the coefficients (pixel j) that appear explicitly

as parameters of the integral; that is, this term is convolutional.

4.3 Use of the FFT to Perform Convolutions in the Internal Field Equations

In Eqs. (3.1e.22) nonzero $\gamma f_\phi(i_x,i_y)$ are available over a region of support $R_{n_{max},n_{max}}$, i.e., $(i_x,i_y)\epsilon R_{n_{max},n_{max}}$ for $0 \leq i_x,i_y \leq n_{max}-1$, and nonzero C_{lj}, or $C(i_x,i_y)$, over a region of support $R_{\infty,\infty}$. It is desired to compute

$$(\gamma f_\phi ** C)(i_x,i_y) \triangleq \sum_{j_x=0}^{n_{max}-1} \sum_{j_y=0}^{n_{max}-1} \gamma f_\phi(j_x,j_y) C(i_x-j_x,\ i_y-j_y) \tag{4.3.1}$$

for the set of Eqs. (3.1e.22)

$$f_{\phi l} = f_{\phi l}^{inc} + \sum_{j=1}^{n_{max}^2} \gamma_j f_{\phi j} C_{lj} \tag{4.3.2}$$

on the region $R_{n_{max},n_{max}}$. One fact that unfortunately can not be made use of is that $C(i_x,i_y)$ is circularly symmetric about the origin (0,0) (namely, because the sampling grid is rectangular). Suppose a function $B(i_x,i_y)$ was convolved with γf_ϕ, where $B(i_x,i_y)$ is of finite support. Let that extent of support be $R_{N,N}$, where for the time being N is chosen to be a little larger than n_{max}. Define $\underline{\gamma f}_\phi(i_x,i_y)$ (read γf_ϕ zero-extended) to be

$$\underline{\gamma f}_\phi(i_x,i_y) = \begin{cases} \gamma f_\phi(i_x,i_y) & (i_x,i_y)\epsilon R_{n_{max},n_{max}} \\ 0 & n_{max} \leq (i_x \text{ or } i_y) \leq N, \text{ the other} \geq 0 \text{ and } \leq n_{max}-1. \end{cases} \tag{4.3.3}$$

Then clearly

$$(\underline{\gamma f_\phi} **B)(i_x,i_y) = \sum_{j_x=0}^{N-1}\sum_{j_y=0}^{N-1} \underline{\gamma f_\phi}(j_x,j_y)B(i_x-j_x,i_y-j_y)$$

$$= \sum_{j_x=0}^{n_{max}-1}\sum_{j_y=0}^{n_{max}-1} \gamma f_\phi(j_x,j_y)B(i_x-j_x,i_y-j_y)$$

$$= (\gamma f_\phi **B)(i_x,i_y). \qquad (4.3.4)$$

If it is now assumed that $\underline{\gamma f_\phi}$ and B are replaced by $\underline{\gamma f_{\phi_p}}$ and B_p, the periodic extensions of the respective finite region of support arrays (sequences), it must be remembered that the result will be shifted repetitions of the aperiodic convolution of $\underline{\gamma f_\phi}$ and B, superimposed. That is,

$$(\underline{\gamma f_\phi}_p **B_p)_p(i_x,i_y) = \sum_{j_x=0}^{N-1}\sum_{j_y=0}^{N-1} \underline{\gamma f_\phi}_p(j_x,j_y)B_p(i_x-j_x,i_y-j_y)$$

$$= \sum_{j_x=0}^{N-1}\sum_{j_y=0}^{N-1} \underline{\gamma f_\phi}(j_x,j_y)B_p(i_x-j_x,i_y-j_y)$$

$$= \sum_{j_x=0}^{N-1}\sum_{j_y=0}^{N-1} \underline{\gamma f_\phi}(j_x,j_y) \sum_{n_x=-\infty}^{\infty}\sum_{n_y=-\infty}^{\infty} B(i_x-j_x-n_xN,\ i_y-j_y-n_yN)$$

$$= \sum_{n_x=-\infty}^{\infty}\sum_{n_y=-\infty}^{\infty} \{\sum_{j_x=0}^{N-1}\sum_{j_y=0}^{N-1} \underline{\gamma f_\phi}(j_x,j_y)B(i_x-j_x-n_xN,\ i_y-j_y-n_yN)\}$$

$$= \sum_{n_x=-\infty}^{\infty}\sum_{n_y=-\infty}^{\infty} (\underline{\gamma f_\phi} **B)(i_x-n_xN,\ i_y-n_yN). \qquad (4.3.5)$$

Thus, it is clear that the convolution of the periodically extended sequences is as described above: a superposition of shifted aperiodic convolutions of $\underline{\gamma f_\phi}$ and B. Note that the period of the repetitions in the output sequence is also N, so the superpositions will overlap. They will overlap because in general the aperiodic convolution of two sequences f_1 and f_2 of regions

of support R_{N_1,N_1} and R_{N_2,N_2} respectively, has region of support $R_{N_1+N_2-1,N_1+N_2-1}$; and so the repeated distributions are larger than the spacing between repetitions. The size of the repeated aperiodic convolutions is determined by examining for what (i_x,i_y) in the aperiodic convolution

$$f_1{**}f_2(i_x,i_y) = \sum_{j_x=0}^{N_1-1} \sum_{j_y=0}^{N_1-1} f_1(j_x,j_y) f_2(i_x-j_x,i_y-j_y) \qquad (4.3.6)$$

is the argument of f_2, $(i_x - j_x, i_y - j_y)$, within R_{N_2,N_2}.

$$0 \le i_x - j_x \le N_2 - 1 \quad \text{or} \quad j_x \le i_x \le N_2 - 1 + j_x. \qquad (4.3.7)$$

The minimum for j_x is zero and the maximum for j_x is N_1 -1. Thus, $0 \le i_x \le N_1 + N_2 - 2$; that is, in the x direction the resulting sequence has length $N_1 + N_2 - 1$. Similar reasoning applies to the y direction. In Fig. (4.3.1) are shown the spatially shifted superpositions of the aperiodic convolution of f_1 and f_2. From the discussion leading to Eq. (4.3.5), the aperiodic convolutions are repeated at spacing N_2. Therefore, repetitions from all sides will interfere over a border $N_1 - 1$ wide; only in the central region will the periodic and aperiodic convolutions be equal. The regions in Fig. (4.3.1) having crosshatch contain overlapped values and are useless in the attempted evaluation of the aperiodic convolution. For the case $N_1 = N_2 = N$ (considered earlier) the region in which the periodic convolution equals the aperiodic convolution shrinks toward zero. The region of support of the result of convolution was shown above to have region of support $N_1 + N_2 - 1$. It is clear that if f_1 and f_2 are zero-padded out to

$R_{N_1+N_2-1,N_1+N_2-1}$ and then periodically extended, yielding $\underline{f}_{1_p}$ and $\underline{f}_{2_p}$, then the resulting convolution ("periodic convolution," which means only that the two sequences being convolved are both periodic) will, within $R_{N_1+N_2-1,N_1+N_2-1}$, equal the aperiodic convolution of f_1 and f_2. Now noting that a periodic sampled function has a discrete Fourier transform and can therefore be written as an inverse transform, one can now write, letting $N \geq N_1 + N_2 - 1$,

$$(\underline{f}_{1_p} ** \underline{f}_{2_p})(i_x, i_y) = \sum_{j_x=0}^{N-1} \sum_{j_y=0}^{N-1} \underline{f}_{1\,p}(j_x, j_y) \underline{f}_{2\,p}(i_x - j_x, i_y - j_y)$$

$$= \sum_{j_x=0}^{N-1} \sum_{j_y=0}^{N-1} \frac{1}{N^2} \sum_{\tilde{i}_x=0}^{N-1} \sum_{\tilde{i}_y=0}^{N-1} \underset{\sim}{f}_{1\,p}(\tilde{i}_x, \tilde{i}_y) e^{j\frac{2\pi}{N}(\tilde{i}_x j_x + \tilde{i}_y j_y)}$$

$$\cdot \frac{1}{N^2} \sum_{\tilde{j}_x=0}^{N-1} \sum_{\tilde{j}_y=0}^{N-1} \underset{\sim}{f}_{2\,p}(\tilde{j}_x, \tilde{j}_y) e^{j\frac{2\pi}{N}\{\tilde{j}_x(i_x - j_x) + \tilde{j}_y(i_y - j_y)\}} \tag{4.3.8}$$

where

$$\underset{\sim}{f}_{1_{2p}}(\tilde{i}_x, \tilde{i}_y) = \sum_{i_x=0}^{N-1} \sum_{i_y=0}^{N-1} \underline{f}_{1_{2p}}(i_x, i_y) e^{-j\frac{2\pi}{N}(i_x \tilde{i}_x + i_y \tilde{i}_y)}. \tag{4.3.9}$$

That is,

$$
\begin{aligned}
&(\underline{f}_{1_p} ** \underline{f}_{2_p})(i_x, i_y) \\
&= \frac{1}{N^2} \sum_{i_{\tilde{x}}=0}^{N-1} \sum_{i_{\tilde{y}}=0}^{N-1} \sum_{j_{\tilde{x}}=0}^{N-1} \sum_{j_{\tilde{y}}=0}^{N-1} [\tilde{\underline{f}}_{1_p}(i_{\tilde{x}}, i_{\tilde{y}}) \tilde{\underline{f}}_{2_p}(j_{\tilde{x}}, j_{\tilde{y}})] e^{j\frac{2\pi}{N}(j_x i_x + j_y i_y)} \\
&= \frac{1}{N^2} \sum_{i_{\tilde{x}}=0}^{N-1} \sum_{i_{\tilde{y}}=0}^{N-1} [\tilde{\underline{f}}_{1_p}(i_{\tilde{x}}, i_{\tilde{y}}) \tilde{\underline{f}}_{2_p}(i_{\tilde{x}}, i_{\tilde{y}})] e^{j\frac{2\pi}{N}(i_{\tilde{x}} i_x + i_{\tilde{y}} i_y)} \\
&\quad \cdot \underbrace{\frac{1}{N^2} \sum_{j_x=0}^{N-1} \sum_{j_y=0}^{N-1} e^{j\frac{2\pi}{N}[j_x(i_{\tilde{x}} - j_{\tilde{x}}) + j_y(i_{\tilde{y}} - j_{\tilde{y}})]}}_{\delta_{i_{\tilde{x}} j_{\tilde{x}}, i_{\tilde{y}} j_{\tilde{y}}}} \\
&= \text{D.F.T.}^{-1}\{\tilde{\underline{f}}_{1_p} \cdot \tilde{\underline{f}}_{2_p}\}.
\end{aligned}
\tag{4.3.10}
$$

This completes the proof and discussion of the two-dimensional discrete convolution theorem. Note that under the conditions of finite region of support for both convolvees, it is essential to zero-pad both functions in order to use the DFT to calculate their aperiodic convolution.

One apparent difficulty in applying the DFT to the convolutional sums encountered in solving the internal field equations for a new estimate of the field is that the region of support of C_{1j}, as noted earlier, is $R_{\infty,\infty}$ (numerically, the C_{1j} are significantly nonzero far outside the extent of the object region). But upon examination of the internal field equations (Eqs. (4.3.2) above), the region of evaluation of the sum is only the object region. Writing these equations in convolutional form (see Section 4.2) results in (leaving implicit in the field arguments the spacing scaling of the indices)

$$f_\phi(i_x,i_y) = f_\phi^{inc}(i_x,i_y) + \sum_{j_x=0}^{n_{max}-1} \sum_{j_y=0}^{n_{max}-1} \gamma f_\phi(j_x,j_y) C(i_x-j_x,i_y-j_y). \tag{4.3.11}$$

Note now that because $0 \leq i_x,i_y \leq n_{max} - 1$ (1 in Eqs. (4.3.2) is a pixel with coordinates $0 \leq i_x,i_y \leq n_{max} -1$), the only values of $C(\cdot,\cdot)$ that contribute to the aperiodic convolutional sum are those for which $-(n_{max}-1) \leq i_x -j_x,i_y -j_y \leq n_{max} -1$. Using the maximum value of j_x, $n_{max} -1$, and the minimum value of i_x, 0, yields the lower limit $-(n_{max} -1)$. Using the minimum value of j_x, 0, and the maximum value of i_x, $n_{max} - 1$, yields the upper limit, $n_{max} -1$. Thus, pictorially, the region of support that contributes to the convolution within the object region is as shown in Fig. (4.3.2). That is, substituting a truncated C as indicated in Fig. (4.3.2) will be mathematically equivalent to "using" the entire C. The solution of the present problem (which the author has been unable to locate in the literature) becomes evident when an alternative, pictorial interpretation of the previous formalized statements is examined. First note that although one has, mathematically equivalently, the convolution of two finite region of support sequences, the C sequence is centered on the origin; that is, the C are nonzero not only in the first quadrant, but in all four quadrants. (Note that in none of the following is the circular symmetry of the C_{lj} recognized or used.) Yet γf_ϕ is centered on $(n_{max}/2 -1,n_{max}/2 -1)$. This problem can not be solved by merely shifting the C sequence, for then the values "called"

in the sum will be incorrect. Consider the pictorial two-dimensional convolution in Fig. (4.3.3a) of two finite region of support sequences f_1 and f_2 centered as would be desired in order to use the DFT. Now periodically extend f_2 after zero-padding it out to N, which is chosen to avoid aliasing, results in $\underline{f}_{2p}$, Fig. (4.3.3b) (where the underline denotes zero-padding and p denotes periodic extension). Only one repetition of f_1 was drawn because, looking at the convolutional sum (Eq. (4.3.8), which can without error be written with sum limits being either N or N_1), it is seen that the argument of f_1 yielding nonzero terms ranges only within R_{N_1,N_1} so that equivalently one need only periodically extend f_2, as far as the evaluation of the convolutional sum, as previously defined, is concerned. (There is no harm in periodically extending f_1 as long as the limits of the spatial area of summation included in the aperiodic convolution are kept in mind.) Consider again the periodic convolution in Fig. (4.3.3b). Another way of viewing the "aliasing" problem is to remember, as noted before, that if, in particular, one is interested in values of the convolution of f_1 and f_2 on R_{N_1,N_1}, the arguments of f_2 will range from $-(N_1 -1)$ to $(N_1 - 1)$. Thus, when "reaching back" for values in, say, the third quadrant, one <u>should</u> obtain zeros. But if f_2 is not zero-padded and then periodically extended, then instead of having Fig. (4.3.3b) one will have the situation depicted in Fig. (4.3.4): now when reaching back for a zero, one will instead obtain a replicated (nonzero) value of $\underline{f}_{2p}$. This is another way of showing why and how much the input sequences must

be zero-padded in order to simulate aperiodic convolution with periodic convolution.

Now consider the contributing block of C for the tomographic convolutions. For the third time, note that these are the only contributing values to the convolution in the region of interest $R_{n_{max},n_{max}}$ If this function is periodically extended, the result is Fig. (4.3.5). Now if one chooses the block of C in the first quadrant region bounded by the axes and the dashed lines as the sequence of finite support to be sent to the FFT, and does <u>not</u> zero-pad, then when evaluating the convolutional sum in $R_{n_{max},n_{max}}$ and "reaching back" into the third quadrant for replications, one obtains (as always) the replica of the upper right-hand side section of the lower left replication of the dashed-line finite sequence, which is <u>exactly equal</u> to what would be obtained in the aperiodic convolution; that is, these very contributions are necessary to simulate the aperiodic convolution. Of course, γf_{ϕ} must be zero-padded out to the size of C, which is chosen for computational convenience to be the lowest power of two greater than or equal to $2n_{max} - 1$. Note that, outside of $R_{n_{max},n_{max}}$, the result will be garbage, but those values can be thrown away.

For an example of this situation consider two cases. The first case is illustrated in Fig. (4.3.6a). In this case, the point of evaluation is inside the object region. Because one replication of C completely covers the object region, the other replications of C make no contribution to the convolutional sum. In the second case (Fig. (4.3.6b)) the point of evaluation is

outside the object region. Note here that other replications of C lie over part of the object region, making contributions that are erroneous with respect to the aperiodic convolution desired. The final result, therefore, upon inverse DFT of $\{\tilde{\gamma f}_\phi \cdot \tilde{C}\}$ where $\tilde{C}$ is the DFT of C as described in the above discussion, is shown in Fig. (4.3.7a), which is the periodic convolution of γf_ϕ with the finite extent, periodically extended C_p. The result in the small square is identical to the result in the same region of the aperiodic convolution of γf_ϕ with the infinite extent C shown in Fig. (4.3.7b).

The above theory is verified practically in Fig. (4.3.8), which shows the real (a) and imaginary (b) parts of the actual convolutions carried out in the spatial (above) and Fourier (below) domains. They agree exactly in the region of interest and the garbage values from the FFT method outside the object region may be discarded. The computational complexity is reduced by this use of the FFT from n_{max}^5 to $n_{max}^3 \cdot \log(n_{max})$.

The digital signal processing technique closest to that just described is the block convolution method called the "overlap-save method." The differences can be reduced to basically two, although the former was not derived from the latter. The first difference is that in the present method, no overlap and no save is performed on the infinite region of support input sequence. This is because only one section of output is desired, which leads to the use of one origin-centered, folded sequence, rather than many unfolded, continuous sections of the input sequence (typically on the positive axis in one dimension rather than two,

for processing of long time signals). The other difference is that in the present application, the region of evaluation of the output is predetermined, and from that the needed section of the infinite region of support input sequence (C_{lj}) is chosen. In overlap-save, a chosen (arbitrary and convenient) sectioning in the infinite region of support input sequence results in valid outputs for each section in a different region determined by the sectioning; thus, the region of support specification reasoning is reversed between the two applications.

Incidentally, if one considers again the possibility of using the FFT by a similar method for the measured scattered field equations (Eqs. (3.1e.21), Section 3.1e), the first consideration will be which values of the D_{mj} contribute to the convolutional sum. Given that the subscript m refers to a receiver and j refers to a pixel in the object region, the difference vectors will fall somewhere on a ring centered a distance R from the origin, given that the receivers are on a ring a distance R from the object region center (see Fig. (4.3.9a)). Consequently, what was a finite square contributing region of support for the C_{lj} (see Fig. (4.3.9b)) becomes roughly a ring contributing region of support, outside of which assuming zero for D_{mj} results in no error. However, such a region is not acceptable for FFT computations, again demonstrating the infeasibility of trying to use the FFT for the scattered field equations convolutions.

4.4 Numerical Implementation

The numerical implementation of the above theory is fairly

direct, except for the generation of the coefficients matrix, which will here be considered. In particular, the basic indexing will be analogous to that in Section 3.2c1c but will have differing ranges. For the spatial domain implementation, the indices for differences in x or y indices ranged from $-(n_{max} - 1)$ to $n_{max} - 1$, but here they will range from $-N'$ to $N' + 1$ where $N' = N/2 - 1$, N being the smallest power of two greater than or equal to $2n_{max} - 1$. Of course, the new range includes the range of importance (the old range). So analogously, the computing order will be

DIFF1	DIFF2
-N′	-N′
"	"
"	"
"	"
-N′	N′+1
-(N′-1)	-(N′-1)
"	"
"	"
"	"
-(N′-1)	N′+1
"	"
"	"
"	"
N′-1	N′-1
N′-1	N′
N′-1	N′+1
N′	N′
N′	N′+1
N′+1	N′+1

(4.4.1)

DIFF 1 and DIFF 2 again are the minimum and maximum differences, respectively, between n - a and m - b (see Section 3.2c1c). Let $\mathcal{N}$ = DIFF 1 + N′ and $\mathcal{M}$ = DIFF 2 + N′. Then there are 2N′ + 2 = N levels above (values of $\mathcal{N}$). The levels, from bottom

up, have 1, 2, 3,... elements. Level $\mathcal{N}$ from the top is N - $\mathcal{N}$ levels from the bottom. The total number of elements in these lower levels is

$$\sum_{i=1}^{N-\mathcal{N}} i = \frac{1}{2}(N-\mathcal{N})(N-\mathcal{N}+1)$$

$$= \frac{1}{2}(N^2+\mathcal{N}^2+N-\mathcal{N})-N\mathcal{N}. \qquad (4.4.2)$$

The total number of distinct coefficients is found by evaluating the above for $\mathcal{N} = 0$:

$$\#\text{DISTINCT COEFFS} = \frac{1}{2}N(N+1). \qquad (4.4.3)$$

Therefore, the total number of elements above in the levels from level zero through level $\mathcal{N}$ - 1 is

$$\frac{1}{2}N(N+1) - \frac{1}{2}(N^2+\mathcal{N}^2+N-\mathcal{N}) + N\mathcal{N} = \mathcal{N}[N+\frac{1}{2}(1-\mathcal{N})]. \qquad (4.4.4)$$

So the index for the beginning of level $\mathcal{N}$ is

$$\mathcal{N}[N+\frac{1}{2}(1-\mathcal{N})] + 1 = \frac{\mathcal{N}}{2}[2N-\mathcal{N}+1]+1. \qquad (4.4.5)$$

From there, to index the correct element within that level, $\mathcal{M}$-$\mathcal{N}$ is simply added (the excess over the value at the beginning of the level). Therefore,

$$\text{INDEX} = \frac{\mathcal{N}}{2}[2N-\mathcal{N}+1]+1+\mathcal{M}-\mathcal{N}. \qquad (4.4.6)$$

To store the portions of the four quadrants in an array that will be transformed by the FFT, for the x (or y) direction the $C_{\Delta x,\Delta y}$ elements are stored as follows: indices 1 through N/2 + 1

receive coefficients with that coordinate difference index having values zero through N/2, and indices N/2 + 2 through N will receive coefficients with that coordinate difference index having values -N′ through -1. Again, to obtain the correct index, first determine DIFF 1 = min (n - a,m - b) and DIFF 2 = max (n -a,m -b). Realizing that, as $\mathcal{N}$ ranges from 0 to N - 1, DIFF 1 ranges from -N′ to N′ + 1, $\mathcal{N}$ = 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. (4.4.6) above. Care had to be used because the grid sizes for Section 3.2c1 and here are generally odd and even, respectively.

In addition to the evidence in Fig. (4.3.8), for an example of a numerical test, a ka = 12.6, 5% speed of sound mismatch cylinder was reconstructed on a 25x25 grid. The appearance of the reconstructions was exactly identical, and the squared errors in $\tilde{\gamma}$ were nearly so:

$$\sum_j (\tilde{\gamma}_j{}^i - \gamma^{ex})^2$$

iteration	spatial domain	FFT
0	2086.2	2086.2
1	306.2	306.2
2	154.8	155.8
3	141.6	141.7
4	140.3	140.2

On the Alliant supercomputer, the execution time using the FFT was 26 seconds, approximately half the time (51 seconds) for the same run (eight iterations) done in the spatial domain.

CHAPTER 5

ADDITIONAL RESULTS

5.0 Introduction

In Chapter 3, the underlying theory of the sinc basis moment method was described in some detail and included numerous results of the present implementation of this method. After a brief discussion of typical tissue characteristics and an abortive attempt to take advantage of the circular symmetry of the C_{1j} coefficients, presented in this chapter will be results of further studies showing dependence of reconstruction quality on method of solution of the nonlinear matrix equations, an unsuccessful proposed alternative solution algorithm, the order of corrections made on the estimated object function, effect of distribution of scattered field energy on correction magnitudes (and a subsequent improvement in convergence speed), and variation of object size for larger objects. Then, more conclusively, degree of phase shift magnitude through the object will be correlated with reconstruction quality for previous reconstruction studies and new contrast experiments for larger cylinders. Plots of these results show a marked threshold of the value of phase shift for successful reconstructions. Then alternative object functions are considered, such as lossy cylinders, a shifted cylinder, and a contrived superposition of simple shapes. Next, the equivalence between the first iteration solution of the sinc basis moment method and that of a representative first-order diffraction tomography algorithm are demonstrated. Finally, a few miscella-

neous but interesting topics are discussed, including the use of incident plane rather than cylindrical waves, the use of line rather than ring geometry for receivers, execution times, and two series of tables of results of the algorithm for various controlled experiments.

Many of the observations in this chapter, particularly those concerned with fundamental limitations of the algorithm such as the phase shift problem and slowness of the reconstruction operations, were elucidated by implementing the programs on an Alliant FX/8 minisupercomputer, located at the University of Illinois Center for Supercomputing Research and Development. This machine has a UNIX operating system, virtual memory, eight 64-bit floating-point processors, fast interprocessor synchronization, and vector instructions. Each processor (Computational Element) can run at 12 megaflops peak (so 96 total). The capacity of memory is 64 megabytes.

5.1 Tissue Characteristics

Typical ranges for speed of sound in human soft tissue are 1.46 mm/μsec (fat) through 1.5 mm/μsec (water) through 1.7 mm/μsec (tendon) and for pressure absorption coefficient from near zero (water) up to 0.04 Np/mm/MHz (abdominal wall) (Goss, et al., 1978). A more typical value for insonification frequency than the 2MHz used in this paper would be 5MHz, which would accentuate the attenuation effects of absorption because of the power law dependence of absorption on frequency. Object size will depend drastically on the particular specimen being examined, but

undoubtedly will be at least hundreds of wavelengths--a long way from the sizes that have been successfully reconstructed in this and other papers. In Section 5.7 and later sections, a limitation of the sinc basis algorithm is identified with regard to phase shift through the object: the magnitude of the phase shift, as defined in Section 5.7, can not exceed π. For a clinically realistic example, a cylinder of diameter 1000 wavelengths and speed of sound 1.7 mm/μsec, the phase shift is about 235 π. Clearly, any practical implementation will have to somehow eliminate this phase shift problem, which exists for many other algorithms as well.

5.2 Circular Symmetry of the C_{lj}

It was hoped that gains could be made in computation time by exploiting the circular symmetry of the C_{lj} coefficients. A program called SPIRAL was developed which counted the number of coefficients having the same radii (in pixel units) and ordered the coefficients such that, in theory, one could take all the terms whose coefficients have a given radius, add them together, and multiply by the appropriate coefficient, saving time and storage space of the C_{lj} (an order of n_{max} fewer coefficients needed). However, the technique proved impractical and far less potentially efficient than using the FFT (Chapter 4).

5.3 Alternative Methods of Solution of the Sinc Basis Tomography Equations

5.3a QR decomposition

This discussion on QR decomposition is a shortened, hybrid form of those in the Linpack manual (Dongarra, 1979) and (Lawson and Hanson, 1974) using the generic notation of this thesis and ammended to enhance the reader's understanding. The QR decomposition of matrix A is

$$Q^T A = \begin{bmatrix} R \\ 0 \end{bmatrix} \tag{5.3a.1}$$

where A is an nxp matrix, Q is an nxn orthogonal matrix, R is pxp, and 0 is (n - p)xp, where $n \geq p$. Because Q is orthogonal, one may write

$$A = Q \begin{bmatrix} R \\ 0 \end{bmatrix} = [Q_1 Q_2] \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 R \tag{5.3a.2}$$

where Q_1 is nxp, Q_2 is nx(n - p), and R is upper triangular--of full rank.

In (Dongarra, 1979), it is now stated, without justification, that $P_A = Q_1 Q_1^T$ and $P_A^{\perp} = Q_2 Q_2^T$ are the projections onto the column space and orthogonal complement, respectively, of A. An argument of justification will here be given. Because Q is orthogonal, its rows and columns are orthonormal vectors: they form an orthonormal basis for $\mathbb{R}^n$. Q_1 must preserve the rank p of R, for A is also of rank p. A column of A is formed by

$$\vec{a}_i = \sum_{k=1}^{p} r_{ki} \vec{q}_{1_k} \tag{5.3a.3}$$

where $\vec{a}_i$ is a column of A, r_{ki} is an element of R, and $\vec{q}_{1_k}$ is a column of Q_1. Thus, the columns of A are linear combinations of the columns of Q_1; in fact the columns of Q_1 span the space of the columns of A. Consider $Q_1Q_1^T$:

$$(Q_1Q_1^T)(Q_1Q_1^T) = Q_1I_pQ_1^T = Q_1Q_1^T. \qquad (5.3a.4)$$

In general,

$$(Q_1Q_1^T)^l = Q_1Q_1^T \qquad l = 1, 2, \ldots \qquad (5.3a.5)$$

which is the definition of a projection. Let $\mathbb{R}^n = V_1 + V_1^{\perp}$ where V_1 is the column space of A and $V_1^{\perp}$ is the orthogonal complement in $\mathbb{R}^n$. Then any vector $\vec{v}$ in $\mathbb{R}^n$ can be written $\vec{v} = \vec{v}_1 + \vec{v}_2$ with $\vec{v}_1 \epsilon V_1$ and $\vec{v}_2 \epsilon V_1^{\perp}$. The columns of Q_1 form an orthonormal basis for V_1. Hence,

$$(Q_1Q_1^T)^l\vec{v} = Q_1Q_1^T\vec{v}_1 \quad \text{and} \quad (Q_2Q_2^T)^l\vec{v} = Q_2Q_2^T\vec{v}_2. \qquad (5.3a.6)$$

So $Q_1Q_1^T$ and $Q_2Q_2^T$ are indeed the respective projections of $\mathbb{R}^n$ onto the space spanned by the columns of A and its orthogonal complement.

To solve the problem $A\vec{x} = \vec{y}$ first write

$$Q^TA\vec{x} = Q^T\vec{y}. \qquad (5.3a.7)$$

If A has rank k (where for the moment $k < p$) let

$$Q^T\vec{y} = \vec{g} = \begin{bmatrix} \vec{g}_1 \\ \vec{g}_2 \end{bmatrix} \qquad (5.3a.8)$$

where $\vec{g}_1$ is a k-vector and $\vec{g}_2$ is an n - k vector.

Then A can be QR-decomposed such that Eq. (5.3a.7) becomes

$$Q^T Q \begin{bmatrix} R_{11} & 0 \\ 0 & 0 \end{bmatrix} \vec{x} \overset{(a)}{=} \vec{g} \quad \text{and} \quad Q^T Q \overset{(b)}{=} I \tag{5.3a.9}$$

where R_{11} is kxk. Decompose $\vec{x}$ into

$$\vec{x} = \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \end{bmatrix} \tag{5.3a.10}$$

where $\vec{x}_1$ is a k-vector and $\vec{x}_2$ is a (p - k)-vector. Then the squared error is, using the fact that orthogonal matrices preserve norms and using Eqs. (5.3a.8) through (5.3a.10),

$$||A\vec{x}-\vec{y}||^2 = ||Q(A\vec{x}-\vec{y})||^2 = ||R_{11}\vec{x}_1-\vec{g}_1||^2 + ||\vec{g}_2||^2 \tag{5.3a.11}$$

and is minimized (and equal to $||\vec{g}_2||^2$) when $R_{11}\vec{x}_1 = \vec{g}_1$. This has a unique solution $\vec{x}_1$ because R_{11} has rank k. So all solutions to the problem of minimizing $||A\vec{x} - \vec{y}||$ are of the form

$$\hat{\vec{x}} = \begin{bmatrix} \hat{\vec{x}}_1 \\ \hat{\vec{x}}_2 \end{bmatrix} \tag{5.3a.12}$$

where $\hat{\vec{x}}_2$ is arbitrary and

$$\hat{\vec{x}}_1 = R_{11}^{-1}\vec{g}_1 = R_{11}^{-1}(Q^T\vec{y})_{1\to k} \tag{5.3a.13}$$

where $(Q^T\vec{y})_{1->k}$ contains the first k elements of $Q^T\vec{y}$, and which equals $Q_1^T\vec{y}$.

Then

$$A\hat{\vec{x}} = QRR_{11}^{-1}Q_1^T\vec{y} = Q_1Q_1^T\vec{y} \tag{5.3a.14}$$

is the projection of $\vec{y}$ onto the column space of A. The residual,

$\vec{r} = \vec{y} - A\vec{x}$, is the projection of $\vec{y}$ onto the orthogonal complement of the column space of A. Therefore, $\vec{r}$ can be found by

$$\vec{r} = Q_2 Q_2^T \vec{y}. \qquad (5.3a.15)$$

In (Dongarra, 1979), $\vec{s}$ is defined as $Q_2^T \vec{y}$. Then $\vec{r} = Q_2 \vec{s}$ (Q_2 twice incorrectly subscripted with a 1 in (Dongarra, 1979)), so the residual norm ρ^2 is

$$\rho^2 = ||\vec{r}||^2 = ||Q_2 \vec{s}||^2 = \vec{s}^T Q_2^T Q_2 \vec{s} = ||\vec{s}||_2^2. \qquad (5.3a.16)$$

Now the effect of this method of solution on the reconstructions is examined. A 17x17 grid was used to reconstruct a ka = 10, 5% speed of sound mismatch cylinder. Figures (5.3a.1) and (5.3a.2) show, respectively, the first iteration speed of sound and absorption reconstructions for four cases. In the upper row of each figure are reconstructions using the routine SQRST from the Linpack software package to solve the scattered field equations for an estimate of the object function, and the bottom rows are corresponding first iteration reconstructions made under the same conditions but calculated via the ART. Define γ^{ex} to be the exact object function and γ^{QR} to be the solution for the object function by QR decomposition. To eliminate one source of trouble, the exact internal field was used for all these reconstructions. In Figs. (5.3a.1a), (5.3a.1c), (5.3a.2a), and (5.3a.2c) this exact field was substituted along with γ^{ex} into the scattered field equations to generate the scattered field data, rather than using the exact series solution as is usually done. Because the

scattered field data were generated by the same equations as are used for reconstruction, the data are a perfect match, and $\gamma^{QR} \simeq \gamma^{ex}$ as shown in Figs. (5.3a.1a) and (5.3a.2a). But if the exact scattered field is used (Figs. (5.3a.1b), (5.3a.1d), (5.3a.2b), and (5.3a.2d)) the solution by QR decomposition is poor; the reconstructed speed of sound in Fig. (5.3a.1b) actually dips below c_0. Using ART with scattered field data generated by the sinc basis scattered field equations, the reconstructions in Figs. (5.3a.1d) and (5.3a.2d) are obtained, which are essentially identical to those obtained for the case of using exact scattered field values (Figs. (5.3a.1c) and (5.3a.2c)). The reason appears to be that in ART one can and does take advantage of the knowledge that $||\gamma||$ is small, and evidently γ^{ex} or, rather, a vector near γ^{ex}, is the smallest local solution of the scattered field equations with exact scattered field data; whereas QR decomposition can make no use of such information. Indeed, γ^{QR} has a norm sixty times greater than that of γ^{ex}. The normalized inner product of γ^{QR} with γ^{ex} is only about 0.3. But not only do Figs. (5.3a.1a) and (5.3a.2a) prove that the routine is "just doing its job," also the same is true for the case of using exact scattered field data (Figs. (5.3a.1b) and (5.3a.2b)). The average percent error in the scattered field for iteration one using ART was as high as 30%, but using QR in this case gave only 0.036% error. Thus, the QR routine is indeed finding the true solution, which is not the one desired but has shifted from γ^{ex} due to the discretization of the integral equation and the expansions over sinc bases.

Later a second investigation of the QR solution was carried out, this time on the supercomputer for a slightly larger object (ka = 12.6, 25x25 grid, 5% speed of sound mismatch). In Fig. (5.3a.3) the QR solution for speed of sound (left) and absorption (right). In the upper row plots the scattered field was generated using scattered field data generated by the sinc basis equations while in the lower row the series solution for the scattered field was used. Again, a perfect match results in nearly perfect reconstruction for the former case (normalized inner product almost exactly one), while for the latter case the results are again poor (average speed of sound within the reconstructed cylinder only 1.4268 mm/μsec compared with the exact value of 1.575 mm/μsec; magnitude of normalized inner product with γ^{ex} only 0.03).

Finally, the other problems with using QR decomposition are the orders of storage (n_{max}^5 compared with n_{max}^3 for ART) and the overall order of computation, which for this algorithm appears to be n_{max}^7 compared with n_{max}^4 for ART using the FFT. For example, on the supercomputer the decomposition time for the 25x25 case took about 25 minutes, and would have to be done many times over a complete iteration (including decomposition of the internal field matrices). When compared with 33 seconds for four complete iterations using ART, there is no contest.

5.3b Alternative ART-type iterative solution

At one time it was proposed that a few points made in (Johnson et al., 1984), be combined into a form that may be suitable

for application of the FFT, as an alternative algorithm using the sinc basis moment method. The algorithm was as follows:

(1) Choose an initial guess for the field (e.g., f_ϕ^{inc}) and for γ (e.g., 0) and therefore for (γf_ϕ) (i.e., 0).

(2) Solve the scattered field equations for $(\tilde{\gamma f}_\phi)^1$ via the ART. Note that n_{trans} smaller problems are solved here (before, all ϕ were combined into one large matrix problem).

(3) Examine the quality of a new estimate of γ given by $\tilde{\gamma}^1$ (evaluated at each pixel).

$$\tilde{\gamma}^1 = \frac{1}{n_{trans}} \sum_{\phi=1}^{n_{trans}} \frac{(\widetilde{\gamma f_\phi})^1}{\tilde{f}_\phi^0} \tag{5.3b.1}$$

(4) For all ϕ perform the convolution of $(\widetilde{\gamma f_\phi})^1$ with C in the Fourier domain (via the FFT) and return to the spatial domain for the new estimate of f_ϕ:

$$\tilde{f}_{\phi l}^1 = f_{\phi l}^{inc} + \sum_{j=1}^{N} (\widetilde{\gamma_j f_{\phi j}})^1 C_{lj} \tag{5.3b.2}$$

(5) Form the new estimate of (γf_ϕ) from $\tilde{\gamma}^1$ of step (3) and $\tilde{f}_\phi^1$ from step (4):

$$(\widetilde{\gamma f_\phi})^2 = \tilde{\gamma}^1 \tilde{f}_\phi^1 \tag{5.3b.3}$$

Use $(\widetilde{\gamma f_\phi})^2$ as the initial guess in step (2)'s use of the ART. Repeat steps (2) through (5) until $\tilde{\gamma}^i$ in step (3) is acceptable.

One new idea in the above algorithm is the solving of the scattered field equations for (γf_ϕ) rather than γ. This should

reduce the computation there by nearly a factor of 2: for each double multiplication necessary before in the set of equations, now only one is necessary. The number of additional operations in step (3) not appearing before is only of the order n^3, and may be even lower than that if it is found unnecessary to average in the terms from all views for the new estimate of γ. Clearly, a substantial increase in memory will occur, nearly double that from before. Perhaps this will make the algorithm unfeasible, perhaps not. But in general it appears that increasing storage by a reasonable amount to save computer time and possibly improve reconstruction quality may be advisable. Note that the order dependence for memory storage is still n^3. Incidentally, if one were to solve the matrix equations using standard methods rather than ART (which requires only one row of the matrix to be in storage at once) that order would be n^5 (see Section 5.3a)! Also new here is the use of the internal field equations to provide a new estimate of f_ϕ rather than to be inverted for f_ϕ.

5.3b1 Later comments

The proposal described in the previous section never came to fruition for the following reasons. What was neglected was that in step two, each of the "smaller problems" had n_{max}^2 unknowns; a total of $n_{trans} \cdot n_{max}^2 \sim n_{max}^3$ unknowns must be calculated given only $n_{det} \cdot n_{trans} \sim n_{max}^2$ measurements. The problem as stated is thus inherently ill-posed. If that were not enough, in step three the field in the denominator can cause trouble at locations where it vanishes or nearly vanishes. This problem is less severe,

because at those locations the contribution from the particular view to the sum for the $\tilde{\gamma}_j$ can be obtained as averages of nearest-neighbors of the $\tilde{\gamma}_j$. However, because the first problem seems fatal, investigation of the proposed alternative algorithm did not extend beyond initial simulations, which were already indicating high sensitivity to noise and poor initial guesses.

As for the statement above about corrections to the field avoiding matrix inversion forms, the correction in Eq. (5.3b.2) is merely that of the simplified ART described in Section 3.3i, with β_0 set to one. As setting β_0 to one (i.e., eliminating it) reduces the number of multiplications and still gives optimum or near-optimum performance for all conditions investigated, this condition was used for most of the runs presented in this chapter.

5.4 Ordering of Corrections to $\tilde{\gamma}$ in the Scattered Field Equations

In the corrections

$$\tilde{\gamma}_j^{i+1} = \tilde{\gamma}_j^{i} + \beta_1 \left(\frac{f_i^{sc} - \vec{r}_i \cdot \tilde{\gamma}^i}{||\vec{r}_i||^2}\right) (\vec{r}_i)_j \qquad (5.4.1)$$

is it important to sequentially modify $\tilde{\gamma}$

or can all the corrections be made on the initial guess, as in

$$\tilde{\gamma}^{l+1} = \tilde{\gamma}^{l} + \beta_1 \sum_{i=1}^{n_{trans} \cdot n_{det}} \left(\frac{f_i^{sc} - \vec{r}_i \cdot \tilde{\gamma}^l}{||\vec{r}_i||^2}\right) \vec{r}_i \qquad (5.4.2)$$

with comparable results (where Eq. (5.4.2) is more amenable to parallel processing than is Eq. (5.4.1))? Computationally, the answer is that the corrections must be made sequentially on a

cumulatively improving $\tilde{\gamma}$. Using Eq. (5.4.1) gave an error of 306.2 after one iteration, compared with an initial error of 2086.2; Eq. (5.4.2) resulted in an error of 9616.4 after one iteration--diverging. A geometrical reason can be given as follows. Consider a 2x2 matrix equation, equivalently the intersection of two lines in a plane (see Section 3.1h). Use $\tilde{\gamma}^0$ as the starting point for $\tilde{\gamma}$ for all projection computations (see Fig. (5.4.1b)) rather than, as usual, using the result of successively corrected $\tilde{\gamma}$s (see Fig. (5.4.1a)) as the starting point for subsequent projection computations. Adding the corrections head-to-tail gives the result of this method--a new $\tilde{\gamma}^1$ which is no closer to γ^{ex} than $\tilde{\gamma}^0$ was, compared with $\tilde{\gamma}^2$ in Fig. (5.4.1a) which <u>is</u> closer to γ^{ex}. Therefore, parallelism can not be taken advantage of in this way to use multiprocessors for reduction of computation time; the ART method is inherently sequential on this level.

Note that the order of corrections makes a difference in convergence time. In Fig. (5.4.2), $\tilde{\gamma}_a{}^2$ is better than (closer to γ^{ex}) than $\tilde{\gamma}_b{}^2$; in general, the estimate $\tilde{\gamma}^i$ should be moved to the hyperplane furthest away from that containing γ^{ex} and $\tilde{\gamma}^i$. The problem is that one can not know these distances beforehand (in Fig. (5.4.2), $|\vec{a}_1|$ and $|\vec{b}_1|$). In general, it is best to order the rows of the matrix equation for maximum orthogonality of adjacent rows. If one examines the elements of the rows for real and imaginary parts of the matrix equations (Eqs. (3.1e.21) and (3.1e.22)) one finds the following structure (verified numerically)

$$\vec{r}_{real} = [a \quad b \quad c \quad d \quad . \quad . \quad .]$$

$$\vec{r}_{imag} = [-b \quad a \quad -d \quad c \quad . \quad . \quad .]. \qquad (5.4.3)$$

It is evident that these adjacent rows are orthogonal;

$$\vec{r}_{real} \cdot \vec{r}_{imag} = 0. \qquad (5.4.4)$$

Therefore, if a reordering of the equations for maximum adjacent orthogonality is to be computed, one should compare only the imaginary part of a current pair with the real part of the candidate next pairs for the inner product minimization search. A result of the recognition of the form of row elements in Eq. (5.4.3) is that a major reduction in computation time resulted (one quarter overall) by computing the elements of only half the rows, and reordering and sign changing to obtain the adjacent row. The results of reordering the rows as described above will be presented in Section 5.14; there it is shown that a huge amount of computation was required to reorder the rows, with a negligible improvement in reconstructions. One reason is that, for pairs, adjacent rows are already exactly orthogonal, and that rearranging pairs will have a relatively minor effect. Also, as $\tilde{\gamma}$ changes during the iterations, the optimal reordering as defined above will change because both $\tilde{\gamma}$ and $\tilde{f}_{\phi}$ are changing, so in theory the rows should be reordered for each iteration, a process that would computationally be hopelessly wasteful and slow. For example, on the Alliant supercomputer, a run that normally took thirty seconds took 27 minutes with this kind of row reordering. As will be shown in the next section, Section 5.5, the overriding priority for row ordering is to make corrections first for regions of largest magnitude of scattered energy.

5.5 Effect of Distribution of Scattered Field Energy

One helpful way of characterizing performance is to examine the variation in absolute value of the error in the estimated scattered field values vs. increasing row correction number i. First, examine Fig. 5.5.1; here is shown the magnitude squared of the scattered field for views 1 and 10 for n_{max} = 25, n_{trans} = 40, n_{det} = 32: a ka = 12.6, c_1 = 1.7391mm/μsec cylinder. Let i_{det} = i modulo (n_{det}). Notice that the peaks occur at approximately $i_{det} = 17 = n_{det}/2 + 1$ and $i_{det} = 24 = [9 \cdot n_{det}/n_{trans}] + n_{det}/2 + 1$ (where [] indicate a greatest integer operation) because that receiver is opposite the transmitter. Also notice how narrow the distribution is about its center. (The slight asymmetries about the center are due to unequal n_{det} and n_{trans} values, causing the sampling angles to be not necessarily symmetrical about the center of the distribution.)

Now consider Fig. (5.5.2): for three values of speed of sound contrast (ka = 12.6) c_1 = 1.7391 mm/μsec (left), 3.00 mm/μsec (center), and 1.4112 mm/μsec (right) [corresponding to phase shifts $\Delta\phi/\pi$ = -1.1, -4.0, and 0.5, (see Section 5.7)], are shown the absolute value of the error in the estimated scattered field, $f^{sc,ex} - \sum_j \tilde{\gamma}_j \tilde{f}^{int}_{D_{mj}}$, for the first 128 rows of iteration 1 (above) and selected rows of iterations 1 through 4 (bottom). Concerning the upper row, notice that the peaks are 32 = n_{det} rows apart, and occur for i_{det} on the point on the ring of transducers opposite the transmitter, where the scattered field energy is maximum (as in Fig. (5.5.1)). The four peaks represent the first four views of correction to $\tilde{\gamma}$. For the first view, the

shape of the peak is close to that of the estimated scattered field distribution shown in Fig. (5.5.1) (though not here squared) because $\tilde{\gamma} \simeq 0$ and the error in the scattered field $\simeq f^{sc}$; it is maximum at $i_{det} = n_{det}/2 + 1 = 17$. For the high contrast case, the shape of f^{sc} is altered, as is f^{sc} in Fig. (5.5.3), due to the increased scattering strength. In the bottom row, the maximum error for each view is plotted: first skip 17 rows, then plot every 32nd row. Thus, for iterations 1 through 4, and $i_{trans} = 1$ through n_{trans}, |error in f^{sc}| is plotted for

$$i = \frac{n_{det}}{2} + i_{trans} + (i_{trans}-1)n_{det} + (i_{iter}-1)n_{det}\cdot n_{trans}. \quad (5.5.1)$$

Note that for $\Delta\psi/\pi \lesssim 1$ (the left and right cases, again see Section 5.7), the error rapidly decreases from a greatly increased value at the beginning of an iteration. There are $n_{trans} = 40$ views (points) per iteration, so iteration 2 begins at 40, iteration 3 at 80, etc. Not so easily explainable is the tendency for the error to rise again near the end of an iteration. That error never really becomes high; perhaps it is just a numerical instability resulting from iterating too long without adjusting the internal field estimate.

Such plots as in Fig. (5.5.2) can help, along with those of squared error in $\tilde{\gamma}$ vs. row of correction in determining the optimal sequence of iterative correction on $\tilde{\gamma}$. Based on the above observations, the behavior in Fig. (5.5.4) can be understood. Figure (5.5.4) is a plot of the squared error in $\tilde{\gamma}$ vs. row of correction for the first 128 rows of iteration 1 of a reconstruction of a ka = 12.6, $c_1 = 1.4118$ mm/μsec, $\Delta\psi = 0.5\pi$

cylinder. Immediately it is evident that the big jumps occur 32 = n_{det} rows apart and coincide with the regions of maximum scattered field energy (because the initial scattered field error is greater, causing greater corrections to be made to $\tilde{\gamma}$). It appears that, by correcting $\tilde{\gamma}$ using first these rows of maximum scattered field energy, the convergence could be speeded up.

Figure (5.5.5) illustrates the advantage of using this idea. The upper row are squared errors in $\tilde{\gamma}$ for a low contrast cylinder (c_1 = 1.4117 mm/μsec, $\Delta\psi/\pi$ = 0.5) and the lower row for a high contrast cylinder (c_1 = 1.7381 mm/μsec, $\Delta\psi/\pi$ = -1.1). The left column shows, under normal row order conditions, the squared error in $\tilde{\gamma}$ after each of the 1280 rows of iteration one. For the low contrast case the error decreases, and for the high contrast case it increases. The small jumps indicate corrections occurring for rows of maximum scattered field energy. The middle column shows, incidentally, the squared error in $\tilde{\gamma}$ for the same points (of maximum error in the scattered field) as given by Eq. (5.5.1). Finally, the rightmost column shows the squared error in $\tilde{\gamma}$ after each of the 1280 rows of iteration one, ordering the rows of correction so that the rows corresponding to the receivers measuring maximum scattered field energy are used first in correcting $\tilde{\gamma}$, and then the others. In this example, defining

$$i_p = 1 + n_{det}\left[\frac{(i_{trans}-1)}{n_{trans}} + \frac{1}{2}\right] \qquad (5.5.2)$$

then the following sequence is used: view by view, beginning with i_{trans} = 1, only receivers i_p - 1 through i_p + 1 are used, for all views; then the other receivers. A dramatic increase in speed

over the leftmost column is evident; about a fivefold decrease occurred in the number of rows it took to reach nearly the final value of $\tilde{\gamma}$ for the iteration. For the general case, a simple, quick determination of the receivers with maximum scattered field energy could be used to implement this idea for a practical application, given that the other problems of practicality of the sinc basis method could be solved. Using this idea, significant underdetermination may be possible without large adverse affects on the reconstruction accuracy. In conclusion, ordering according to locations of largest scattered field energy is far more effective (and determination of order far faster) than reordering according to minimization of adjacent row inner products (given that real and imaginary part rows, which have always been kept adjacent, are already orthogonal).

5.6 Object Size for Large Objects

In this section, reconstructions of lossless, 5% speed of sound mismatch circular cylinders are presented for varied object size, ranging from ka = 18 to ka = 32. Three-dimensional views of the first iteration reconstructions of speed of sound (Fig. (5.6.1)) and absorption (Fig. (5.6.2)) illustrate the declining behavior of the Born approximation with object size. Also noticeable from these plots is the smoothness and the circular symmetry of the first iteration reconstructions, and a general following of the cylinder boundary as it gradually grows (a pixel at a time from plot to plot) from left to right, top to bottom of the composite. A certain amount of mixing of the real and imaginary

parts of the object function is evident, for the absorption reconstructions obviously have nonzero average values. Also, the speed of sound reconstructions droop badly within the cylinder increasingly as the cylinders grow in size.

Turning to the fourth iteration, Figs. (5.6.3) and (5.6.4) present, respectively, the speed of sound and absorption reconstructions. Here even in the perspective views, which are not quantitatively comparable with the exact solution, the failure for larger objects is obvious. For small objects, reconstructions are quite good, the error for both speed of sound and absorption reconstructions being significant only at the cylinder boundary. But the circular symmetry is obviously lost for larger objects until, for ka = 32, the reconstruction is hardly recognizable as a lossless circular cylinder.

Iterations one and four are summarized and quantitatively compared with the exact solution in Figs. (5.6.5) and (5.6.6), which are, again, respectively speed of sound and absorption composites. One can see that even for small objects the first iteration (Born approximation) yields highly erroneous speed of sound and absorption values within the cylinder, while the fourth iteration is a dramatic improvement, oscillating closely about the exact solution values. But as the size increases to about ka = 25, even the sinc basis method begins to fail and become asymmetrical, as was seen also in Figs. (5.6.3) and (5.6.4). This effect appears to be a numerical artifact of this implementation of the algorithm, occurring when conditions for the validity of the equations are only marginally satisfied. Because only the

measured scattered field equations are used to generate the first iteration reconstruction, the symmetry present for all sizes isolates the origin of the problem as being in the other set of equations, the internal field equations. The position of the peak in the asymmetry has been shown to depend on the order of the viewing angles in the iterative correction within the internal field equations. It was thought that because most of the correction is done during whichever view is first considered, by underrelaxing the first few views relative to the others this effect might be reduced. However, for the first three views, changing the relaxation constant from 1.0 to 0.2 changed the detail in but not the degree of asymmetry. Figure (5.6.7) shows the results of this study, performed on a ka = 12.6, -11% speed of sound mismatch lossless cylinder. Iterations 5 and 8 are plotted against the exact solution, and in nearly all cases exhibit asymmetry. For $\beta_0 = 0.2$, the asymmetry does not appear until somewhere between iterations 5 and 8, but it appears nonetheless, indicating that overcorrection is not the source of the problem although underrelaxation may to some degree slow down asymmetry generation (as well as the reconstruction process in general).

In further iterations of the object size study the asymmetry effect is altered and reduced, but not eliminated. Iteration 8 is shown in Figs. (5.6.8) and (5.6.9). Compared with iteration four, the reconstructions are somewhat smoother, and the Gibbs-type phenomenon on the cylinder boundary is reduced or nearly eliminated. Evaluation can be refined by examining the center-line profiles of iterations 5 and 8 and the exact solution, shown in

Figs. (5.6.10) and (5.6.11). From ka = 23.3 through 26.8 (second row of the composite) there is substantial movement in both speed of sound and absorption reconstructions toward the exact solution. But for ka = 28.6 and beyond, improvements diminish and the erroneous asymmetry dominates the appearance of the images. Further studies could assess the effects of other system parameters on symmetry of reconstruction. This problem will be reconsidered shortly. It is thus seen how critical the accuracy of the estimated internal field is for successful reconstruction; any systematic errors there will severely distort the reconstructed object function.

5.7 Phase Shift Through the Object

At this point, it is helpful to define what is meant by "phase shift through the object." If a wave undergoes a phase change from ψ_0 to ψ_1 when travelling a distance R through free space (see Fig. (5.7.1)), and undergoes a phase change from ψ_0 to ψ_2 when travelling the same distance, but in its course also passing through the diameter of a circular cylinder of different speed of sound, the difference in phase shifts is called $\Delta\psi$. That is,

$$\Delta\psi = \psi_2 - \psi_1 = \omega(2a)\left(\frac{1}{c_1} - \frac{1}{c_0}\right). \qquad (5.7.1)$$

The basic hypothesis is that when $|\Delta\psi|$ exceeds π, there will be information lost about it when sines and cosines are taken for the representation of the field during the reconstruction compu-

tations, resulting in a nonuniqueness of solution. Essentially, $2\pi + \pi/4$ becomes indistinguishable from $\pi/4$.

5.7a Phase shifts of cylinders reconstructed in Chapter 3

In Fig. (3.3e.2) of Section 3.3e, the early study examining contrast variation for three small object sizes, the only case for which higher iterations were degradations rather than improvements over the lower iterations was that for ka = 13.6, -10% contrast. It is also the only case in that entire composite for which $|\Delta\psi| > \pi$ ($\Delta\psi = 1.11\pi$).

The largest cylinder in the early object size study (Section 3.3f), ka = 18.9, with -5% speed of sound contrast, was successfully reconstructed, but its phase shift was only 0.63π.

5.7b Phase shifts of cylinders in object size study of Section 5.6

In the object size study of Section 5.6 the phase shift ranged from $-\pi/2$ to $-\pi$. There one finds that, around $\Delta\psi = -0.87\pi$ (beginning of the third row) reconstruction quality is already poor. In fact it will be seen in Section 5.7d that often the sinc basis method fails slightly before $|\Delta\psi| = \pi$ is reached. This is probably due to the discretization and sinc basis expansion limitations on the representation of the various fields. However, the results of that object size study could have been made to appear somewhat more favorable and uniform by increasing the sampling density with decreasing cylinder size to maintain a constant sampling density to estimated Nyquist density but, because in

general the object function is not known ahead of time, such an action seems unjustifiable. For this reason, the other of the two ways of varying the phase shift--varying the speed of sound contrast as opposed to object size (see Eq. (5.7.1))--produces results that are more easily and closely comparable. (Indeed, results for reconstructing very small cylinders on large grids where h was chosen for inclusion of the spatial frequency content of a large cylinder were markedly worse than the reconstructions of the same object on a smaller grid with smaller spacing; if the spacing of the large grid was made as small as that chosen for the smaller grid, results became identical.)

5.7c Contrast studies for larger (ka = 32 and 46) cylinders

Figures (5.7c.1) through (5.7c.4) are reconstructions of ka = 32 cylinders for varying speed of sound contrast. Speed of sound (left) and absorption (right) reconstructions are plotted against the exact solution for decreasing positive contrast (Fig. (5.7c.1)) for iterations one and four, and increasing negative contrast (Fig. (5.7c.2)). Figures (5.7c.3) and (5.7c.4) show the same reconstructions for iterations five and eight. The phase shift ranges from -1.42π to 1.46π. One can see a maximum of asymmetry for the cases $\Delta\psi \simeq \pm\pi$ (c_1 = 1.575 mm/μsec and 1.425 mm/μsec) along with a minimum of stabilization (e.g., from iteration 5 to 8) relative to the reconstructions for other contrast values. For the low contrast cases (+1%, $\Delta\psi \simeq 0.2\pi$) the reconstructions for this larger size cylinder are highly accurate in shape and value. For the very large contrast cases (7.5% and

-6.7% contrast, $\Delta\psi = -1.42\pi$ and 1.46π) there begins to appear a stabilizing erroneous solution somewhat akin to but clearly distinct from the Born approximation solution (i.e., iteration one). The contrast was not taken higher still for this early composite. A subsequent study, however, did examine higher contrast cases for a yet larger object: ka = 46.

In Figs. (5.7c.5) through (5.7c.8) are shown 75x75 reconstructions of ka = 46 cylinders of varying speed of sound contrasts ranging from 7.5% to -6.7% ($\Delta\psi$ from -2π to 2π). Again note the very high quality of reconstructions for low contrast cases, and a tendency towards asymmetry and loss of accuracy in intermediate contrast cases. But notice that for very high contrast cases, where for example the phase shift is $\pm 2\pi$, the reconstructions again become symmetrical. The reconstruction does not blow up, oscillate, or randomly wander, but settles on an erroneous solution. That supports the hypothesis that above π, there is a nonuniqueness phenomenon taking the iterated solution away from that desired and along the path of a solution with phase shift magnitude less than π. The phase shifts through the reconstructed cylinders were calculated as follows:

$$\widetilde{\Delta\psi}^j = \omega h \sum_{i=1}^{n_{max}} \left(\frac{1}{\tilde{c}_i^j} - \frac{1}{c_0}\right) \qquad (5.7c.1)$$

where $\widetilde{\Delta\psi}^j$ is the jth iteration phase shift through a midline of the object region and $\tilde{c}_i^j$ is the speed of sound estimate for pixel $(i_x, i_y) = (i, n_{max}/2 + 1)$ for iteration j. Almost invariably the phase shifts were of magnitude less than π. Perhaps it is

this retracking of the solution that may account for the subtle artifact of asymmetry that occurs only in the vicinity of $\pm\pi$.

As an aside, Fig. (5.7c.9) shows an example of a 101x101 (ka = 63.6) reconstruction for a 1% speed of sound mismatch. The upper row are speed of sound and absorption reconstructions for iterations one and four plotted against the exact solution, while the bottom row is the same for iterations five and eight. The high accuracy is very encouraging, but the computation time was very long (three hours and 12 minutes) and furthermore the program aborted even for one case where $|\Delta\psi| < \pi$. Such behavior may be attributed to a machine problem, but further attempts were not made because the Alliant computer became for this large problem overloaded and in a state of thrashing (the computer comes to a standstill for all users).

The behavior described above with respect to phase shifts is summarized concisely in Fig. (5.7c.10). Shown there are speed of sound reconstructions for a ka = 32 cylinder for three contrast values: -2.5%, -5%, and -17% with corresponding phase shifts $\pi/2$, π, and 4.2π. For the low contrast object the fourth iteration shows dramatic improvement over the first iteration reconstruction. Significant asymmetry becomes evident for the medium contrast object (in the vicinity of π), and symmetry returns for high contrasts but the (stable) solution is erroneous, with reconstructed phase shift of magnitude less than π.

One might ask if the reconstructed phase shift in some way follows the correct phase shift, modulo π. This possibility was investigated in Table (5.7c.1) for ka = 12.6 cylinders. It is

seen that there is no discernible pattern even though for a few cases ($\Delta\psi/\pi$ = 4.2, 2.1 and 2.7) there seems to be a strong correlation with $\widetilde{\Delta\psi} = \Delta\psi$ modulo(π). Because of the other cases, however, these must be viewed as merely coincidences. It must here be noted that the situation is more complicated than a characterization of the sinc basis method as being simply a means of reconstructing phase shifts (e.g., many signals of erroneous phase shifts are combined together for the iterative corrections), though the reconstructed phase shift is one criterion for judging the accuracy of reconstructions. (For comparison, two relatively low contrast cases are given in Table (5.7c.1): $\Delta\psi/\pi$ = 0.5 and 0.8. After eight iterations the phase shifts are off by only 1.6% and 1.5%., respectively.)

5.7d Simulation results plotted against $\Delta\psi/\pi$

This section will present and analyze a series of plots of various errors and improvement factors. The horizontal axis for all plots is the phase shift through the diameter of the lossless cylinder divided by π. The four curves are results of reconstructions for the fixed object sizes ka = 4.7, 12.6, 32, and 46, corresponding to objects of radii 0.75, 2.0, 5.1, and 7.3 wavelengths. Each point on each curve represents the error or improvement factor for an individual run under conditions of the given object size and speed of sound contrast (phase shift). The first plot, Fig. (5.7d.1), is the log of the squared error normalized to the number of pixels in the cylinder after zero iterations. A given phase shift requires larger contrast values and

therefore larger $|\gamma|$ for smaller cylinder radii (see Eq. (5.7.1), Section 5.7). The initial error is in fact the exact object function because the initial guess is zero; that is,

$$E_0 = \log_{10}\left(\frac{||0-\gamma^{ex}||^2}{n_{\gamma\neq 0}}\right) = \log_{10}\left(n_{\gamma\neq 0} \cdot \frac{(\gamma_{cyl}^{ex})^2}{n_{\gamma\neq 0}}\right)$$

$$= 2\log_{10}(\gamma_{cyl}^{ex}) \qquad (5.7d.1)$$

where γ_{cyl}^{ex} is the value of γ^{ex} within the cylinder. Thus, the initial error will be higher for decreasing cylinder radii, and so the four curves are separate. The same minimum contrast value was chosen for all sizes: ±1%, so the normalized initial error is the same for all four sizes (and note that the phase shift for ±1% increases with object size). The minimum value of the curves follows from the value of γ for a 1% contrast cylinder:

$$\gamma_{1\%} = (2\pi \cdot 2)^2 \left(\frac{1}{1.515^2} - \frac{1}{1.5^2}\right) = -1.384 \text{ (rad/mm)}^2 \qquad (5.7d.2)$$

so $E_{min} = 0.282$ (> 0).

The next plot, Fig. (5.7d.2), is the same error after the first iteration, which is based on the Born approximation. Nearly the same relation between the curves is maintained as in the initial error plot (Fig. (5.7d.1)). Improvement over the initial guess is limited to the cases where $|\Delta\psi| << \pi$, where E drops below zero (which here merely indicates that $\sum_j(\tilde{\gamma}_j^1 - \gamma_j^{ex})^2 < n_{\gamma\neq 0}$). Outside $\pm\pi$ the error is no lower than that of the initial guess.

After four iterations (Fig. (5.7d.3)) one can see dramatic

improvements in the error vs. phase shift within $\pm\pi$; for a given phase shift the error curve is pulled down. In fact, the amount pulled down increases somewhat with object size, for intermediate sizes. However, the few 101x101 results obtained for four iterations are added as separate points to the plot in Fig. (5.7d.3), and, together with the 55x55 and 75x75 curves show that this trend does not continue for larger objects. Again, outside of $\pm\pi$ the error is no lower than that of the initial guess.

Figure (5.7d.4) shows the error after eight iterations. Here there is small but definite improvement over the fourth iteration. (One can place Fig. (5.7d.3) over Fig. (5.7d.4) to see this; the gains are greatest in the vicinities of $\pm 0.6\pi$ through $\pm 0.8\pi$.) In addition, shown as separate points are the results of the object size study of Section 5.6. Each of those points is a 5% contrast cylinder of radius 17 through 10 pixels reconstruction (left to right). As could be expected, the smaller sized cylinders have decreasing phase shift magnitudes for the fixed contrast, and consequently result in lower errors. The slight upturn in the error for the smallest few cylinders is due to the point referred to in Section 5.7b--the problem of incomparability of results when using the same grid spacing for different problems. In an experimental setup, some estimate would have to be used for the minimum object size likely to be encountered, in choosing the grid spacing. In addition, in Fig. (5.7d.4) the only two runs for which the 101x101 program completed eight iterations (±1%) are plotted as separate points, and are found just below the 75x75 curve. They again indicate that the situation, even

within $\pm\pi$, will not indefinitely improve with increasing object size, even neglecting the enormously growing computational loads.

Figure (5.7d.5) is a set of plots of $10\log_{10}\{||0-\gamma^{ex}||^2/||\tilde{\gamma}^1-\gamma^{ex}||^2\}$, the dB improvement in the squared error vs. $\Delta\psi/\pi$ after 1 iteration (the Born approximation) over the initial error. Note the relatively low dependence on object size--the curves are all close together. The closeness indicates that this one basic curve may represent a fundamental characteristic of the result of using the first Born approximation for tomographic reconstruction. There are improvements of up to 10 dB for very low contrast (1%), but outside of $\pm 0.85\pi$ the dipping below zero indicates a small degradation over the initial guess for γ of zero. The value of this degradation, interestingly, remains fairly constant for large ranges of $\Delta\psi$ outside $\pm\pi$.

In Fig. (5.7d.6) is shown the improvement factor in dB after 4 iterations over the initial guess vs. phase shift increment. The encouraging feature here is that the improvement increases with object size, and is relatively flat within $\pm\pi$, peaking at near 20 dB, compared with 10 dB for the Born approximation. Again, because of the nonuniqueness problem occurring for phase shifts greater than π, the algorithm fails outside the region where the "minimum norm" solution as obtained by ART (see Section 3.1i) coincides with the desired solution.

The improvement factor after eight iterations over the initial guess is plotted in Fig. (5.7d.7). The main improvement here over that in Fig. (5.7d.7) is the increased flatness of the curves within $\pm\pi$. To show that the curves are indeed flat over

many separate runs, the individual run results are for this plot indicated by the symbols printed on top of the curves. The flatness of the curves shows that when the assumptions of the algorithm are valid, the sinc basis method yields a solution whose improvement over the initial guess of zero is fairly independent of the phase shift through the diameter of the cylinder. Here, as $\pm\pi$ are approached from zero, additional limitations are found on performance as $\pm 0.85\pi$ (roughly) are reached. This phenomenon of beginning to fail slightly before $\pm\pi$ presumably results primarily from the poor initial guess for the internal field (the incident field), and possibly also from the discretization of the original integral equations, the sinc basis field representation, and other numerical errors.

It is interesting to note that while the Born approximation gradually deteriorates to approximately no improvement at $\pm$ 0.8π (see Fig. (5.7d.5)), the sinc basis method fails abruptly at approximately $\pm\pi$. This behavior indicates that the phase shift nonuniqueness problem is now the "weakest link" in the algorithm.

Figure (5.7d.8) shows the improvement factor after 4 iterations over the first iteration (Born approximation) vs. phase shift:

$$S_{14} = 10\log_{10}\left(\frac{||\tilde{\gamma}^1 - \gamma^{ex}||^2}{||\tilde{\gamma}^4 - \gamma^{ex}||^2}\right). \qquad (5.7d.3)$$

Because of the property of the log of a quotient, Fig. (5.7d.8) represents the differences between the curves in Fig. (5.7d.6) and (5.7d.5). For very small contrasts the Born approximation is already good, so the improvement factor is small. The improvement

factor steadily increases for increasing phase shifts to a peak of about 16 dB until $\pm\pi$ are approached, at which point the algorithm begins to fail. This increase is due to the increasingly poor Born solution occurring, while the improvement of the sinc basis method is relatively constant within this range of $\Delta\psi$. If the solution to the difficult problem of adding signals with unwrapped phases were known and could be applied here it is conceivable that the improvement factor could continue its increasing trend outside the $-\pi$ to π interval.

Figure (5.7d.9) shows the improvement factor after 8 iterations over the first iteration (Born approximation) vs. phase shift. Essentially the same patterns hold as for Fig. (5.7d.8); only the maximum improvement increases from around 16 dB to about 18 dB.

Figure (5.7d.10) gives the absolute value of the percent error in the initial estimated average speed of sound (which is that of the homogeneous coupling medium, c_0) within the cylinder. That is, it shows $100|(c_0 - c_1)/c_1|$. The curves are straight lines because $\Delta\psi/\pi = 4fa\cdot(1/c_1 - 1/c_0)$ (f is the frequency) is linearly related (within an additive constant) to the inverse of the actual speed of sound in the cylinder c_1. For a given phase shift, again the percent error decreases with increasing object size because a lower speed of sound contrast (and therefore a lower magnitude error in the initial guess c_0) is required to obtain the given phase shift.

Figure (5.7d.11) plots $100|(\tilde{\bar{c}}_1^1 - c_1)/c_1|$, the absolute value of the percent error of the Born approximation-based aver-

age speed of sound estimate within the cylinder. The improvement within $\pm\pi$ over the initial guess is indicated by noting that generally these percent error curves dip below the straight lines of the previous plot, increasingly for lower contrasts. (Again, one can place Fig. (5.7d.10) over Fig. (5.7d.11) to compare the plots, because the axes scaling happens to be the same for the two.)

The absolute value of the percent error of the 4th iteration average speed of sound estimate within the cylinder, $100|(\tilde{c}_1{}^4 - c_1)/c_1|$, is shown in Fig. (5.7d.12). The improvement within $\pm\pi$ over both the initial guess and that of the Born approximation is dramatic. Further encouraging is the fact that the lowest curve is that for the largest object. (However, note also that for these cases the percent error in the initial estimate for c_1, c_0, is lower for larger objects for a given phase shift value.) For the largest object, ka = 46,

1% contrast <=> $\Delta\psi/\pi$ = -0.3 -> -0.014% error

-3.3/3.5% contrast <=> $\Delta\psi/\pi = \pm 1.0$ -> 1.6/-1.7% error

$\pm$ 2.5% contrast <=> $\Delta\psi/\pi$ = .75/-.7 -> -.005/.012% error

(A note on the slight asymmetries: for a given phase shift magnitude, the magnitudes of the object function for positive and negative phase shifts differ because of the asymmetrical dependence of the object function on c_1 and therefore on the phase shift. In any case, the computations involved in computing the scattered field and reconstruction yield complicated, highly nonlinear functions relating phase shift and percent contrast. Outside $\pm\pi$, the percent errors quickly become very large.)

Figure (5.7d.13) gives $100|(\tilde{c}_1^8 - c_1)/c_1|$, that is, the absolute value of the eighth iteration percent error in the estimation of the average speed of sound within the cylinder. The results are much the same here as in Fig. (5.7d.12). For the largest (ka = 46) cylinder

1% contrast <=> $\Delta\psi/\pi$ = -0.3 -> -0.009% error

-3.3/3.5% contrast <=> $\Delta\psi/\pi$ = ±1.0 -> 1.4/-1.1% error

±2.5% contrast <=> $\Delta\psi/\pi$ = 0.75/-0.7 -> 0.06/-0.03% error

all of which are small improvements over or comparable with the corresponding errors after four iterations. But again, outside $\pm\pi$, the algorithm breaks down and the percent errors in c_1 skyrocket.

With regard to this failure, it has been suggested (private communication, M. Soumekh) that as here only complex field operations and no transformations of the phase such as that for the complex log are used in this formulation, that even if it were possible to account for total phase, it would be of no advantage here. In any case, a complete understanding of this problem requires further study. For example, in light of the work in (Berggren et al., 1987) where claims of (slow) convergence at or beyond π using the sinc basis moment method were made, the negative conclusions given here must be viewed as tentative. Other methods of solution and additional knowledge of various characteristics of the field and/or the object function might be incorporatable into more advanced implementations.

For a final set of percent error plots, Figs. (5.7d.14) through (5.7d.17) present the same errors as in Figs. (5.7d.10)

through (5.7d.13), but are normalized to the exact percent contrast. This was done to remove the effect of the fact mentioned in the discussion of Fig. (5.7d.12) that for increasing object size, a given phase shift will have decreasing initial error in speed of sound estimation. So, one would expect that by normalizing to the percent contrast one could at least partially uniformize the curves, and better represent a fundamental limitation of the present implementation of the sinc basis moment method. Fig. (5.7d.14) gives the absolute value of the initial percent error in the estimated speed of sound within the cylinder, divided by the percent contrast:

$$\left|\left(\frac{c_0-c_1}{c_1}\right) \Big/ \left(\frac{c_1-c_0}{c_0}\right)\right| = \left|-\frac{c_0}{c_1}\right| \qquad (5.7d.4)$$

which is simply a set of linear curves. The curves intersect at $\Delta\psi = 0$, normalized error = 1, and the angle with respect to the horizontal increases with decreasing cylinder size, for the reason just explained.

Figure (5.7d.15) gives the same error after one iteration:

$$\left|\left(\frac{\tilde{c}_1-c_1}{c_1}\right) \Big/ \left(\frac{c_1-c_0}{c_0}\right)\right| = \left|\frac{c_0(\tilde{c}_1-c_1)}{c_1(c_1-c_0)}\right| \qquad (5.7d.5)$$

Here the reason for dividing by percent contrast is evident. In Fig. (5.7d.11), the curves for different sizes are widely spread out because of the fact about percent contrasts vs. phase shift vs. size, given above. In Fig. (5.7d.15) the curves are all much closer together, indicating a general behavior of the algorithm rather than merely various results for particular objects.

Figure (5.7d.16) gives

$$\left|\left(\frac{\tilde{c}_1^4-c_1}{c_1}\right) \Big/ \left(\frac{c_1-c_0}{c_0}\right)\right| = \left|\frac{c_0(\tilde{c}_1^4-c_1)}{c_1(c_1-c_0)}\right| \qquad (5.7d.6)$$

For low contrasts the curves are very similar to those in Fig. (5.7d.12) (the corresponding unnormalized percent error plot), but for high contrasts, especially outside $\pm\pi$, the curves are brought much closer together in Fig. (5.7d.16).

Finally, in Fig. (5.7d.17) the same error for iteration eight is plotted, which is the normalized version of Fig. (5.7d.13). The uniformity is even more marked here and in Fig. (5.7d.16) if the curve for the smallest cylinder (ka = 0.75) is removed, which for all cases is the curve off by itself (probably due to a high discretization error caused by a small number of pixels representing the circular shape of the object).

5.7e Attempting to add signals with unwrapped phase

In Section 5.7d it was suggested that if signals of unwrapped phase could be reliably and uniquely added, there might exist a way to make use of such absolute phase information to extend the region of high performance (such as that shown within $\pm\pi$ in Figs. (5.7d.8) and (5.7d.13), for example) outside the $(-\pi,\pi)$ interval. Thus, this subsection addresses the problem of adding two signals of unwrapped phase. The first part of this discussion is based on a talk with M. Haney.

Given two signals with unwrapped phases, $Ae^{j\varphi_A}$ and $Be^{j\varphi_B}$, what is their sum $Ce^{j\varphi_C}$ (its magnitude C and unwrapped phase φ_C)?

In particular, the phases are $\varphi_A = n_A 2\pi + \phi_A$, $\varphi_B = n_B 2\pi + \phi_B$, and $\varphi_C = n_C 2\pi + \phi_C$ where n_A, n_B, and n_C are integers and ϕ_A, ϕ_B, and ϕ_C are real with absolute value less than 2π. For example, suppose that two identical, coherent point sources radiate through regions of differing speed of sound and then re-enter a homogeneous space. How would the two signals combine at point P [see Fig. (5.7e.1)]? The following phasor diagrams illustrate the difficulty in obtaining a unique value of the phase when one or both of n_A and n_B are nonzero. Suppose $B < A$ and $0 \leq \varphi_A < 2\pi$. Then in Fig. (5.7e.2a) the resulting phase, as intuitively expected, will be between ϕ_1 and ϕ_2. If now $B > A$ the head-to-tail figure is as shown in Fig. (5.7e.2b) and for $B = A$ the rotations pass through the origin. In these latter cases it would appear that as the winding number of signal B increases, so will that of signal C, contrary to the case $B < A$. Plotting the phase vs. rotations of the B signal for the two situations yields Fig. (5.7e.3), from which it is seen that as B approaches A, a unique phase is not obtained; that is, phase(B↑A) ≠ phase(B↓A).

If the sum is written as

$$e^{j\varphi_A}(A + jBe^{j(\varphi_B - \varphi_A)}] \qquad (5.7e.1)$$

it is easily seen that, depending on whether $B < A$ or $B > A$ the phasor will cross the origin, thereby increasing the winding number (or, the phase plane from, e.g., $0 \to 2\pi$ to $2\pi \to 4\pi$, etc.), or whether the phasor will rotate within the same plane (e.g., $0 \to 2\pi$).

A second way to consider the problem is to write the series expansion for the exponential function

$$Ae^{j\varphi_A} + Be^{j\varphi_B} = A\sum_{n=0}^{\infty}\frac{(j\varphi_A)^n}{n!} + B\sum_{n=0}^{\infty}\frac{(j\varphi_B)^n}{n!}$$

$$= \sum_{n=0}^{\infty}\frac{j^n}{n!}(A\varphi_A^n + B\varphi_B^n)$$

$$= X + jY \quad \text{where} \begin{cases} X = \sum_{n_{even}}\frac{j^n}{n}(A\varphi_A^n + B\varphi_B^n) \\ Y = \sum_{n_{odd}}\frac{j^{n-1}}{n}(A\varphi_A^n + B\varphi_B^n) \end{cases}$$

$$= Ce^{j\varphi_C}$$

$$= \sum_{n=0}^{\infty}\frac{j^n}{n!}C\varphi_C^n = \sum_{n_{even}}C\frac{(j\varphi_C)^n}{n!} + j\sum_{n_{odd}}Cj^{n-1}\frac{\varphi_C^n}{n!}$$

(5.7e.2)

The problem here is that only the sums need be equal and not the individual terms, which makes solving for C and φ_C impossible. Difficulties such as encountered here make use of the unwrapped phase difficult if not impossible in the sinc basis moment method. The problem of possible calculation and use of unwrapped phase of sums of unwrapped phase signals (which is exactly what in theory takes place in Eqs. (3.1e.21) and (3.1e.22)) is left as an open area for further study.

5.7f Use of alternative initial guesses for γ and f^{int}

5.7f1 Initial guess for γ is zero

In order to find out more about the nature of local solutions obtainable by ART with an initial guess of zero for γ and

about the accuracy requirements for the field, a 25x25 speed of sound contrast study of lossless, ka = 12.6 circular cylinder reconstructions was performed that 1) compared reconstruction quality for three field conditions as outlined below and 2) used symmetrical about zero values of $\Delta\psi/\pi$ as the independent variable. Pairs of figures, Fig. (5.7f1.1) and (5.7f1.2), (5.7f1.3) and (5.7f1.4), and (5.7f1.5) and (5.7f1.6) show, respectively, negative and positive phase shift speed of sound reconstructions for iterations 1 and 4, 5 and 8, and 9 and 12, all plotted against the exact solution. In the first (leftmost) and second (middle) columns of each figure, respectively, the incident field and exact internal field as obtained in Section 3.2d1 is used as the initial guess for the internal field, and both Eqs. (3.1e.21) and (3.1e.22) are used, as usual, to solve for better estimates of both γ and f^{int}. In the third column, the exact internal field is used as the internal field estimate for <u>all</u> iterations (no iterative corrections on the field are performed using Eqs. (3.1e.22)).

First consider reconstructions in the first (leftmost) column. As expected from previous studies, when using the incident field as the initial guess for the total internal field, the algorithm succeeds for low contrast cases ($\Delta\psi/\pi = \pm 0.5$ and ± 0.8), but breaks down for phase shift magnitude greater than π. Note, however, in iterations 9 and 12 (Figs. (5.7f1.5) and (5.7f1.6)) that the $\pm\pi$ reconstructions are on the threshold of being acceptable. Note also that in the vicinity of π ($\pm\pi$, $\pm 1.1\pi$, and $\pm 1.2\pi$) that these reconstructions are most highly asymmetri-

cal, and changing with iteration even after 9 iterations. Very large phase shifts ($\pm 2\pi$ and $\pm 4\pi$) show a return to stability, but on an erroneous solution (with phase shift magnitude less than π).

Now consider the middle column for Figs. (5.7f1.1) through (5.7f1.6). Use of the exact internal field as the initial guess for the internal field had the following effects. Within $\pm\pi$ the final reconstructions are about the same as those using the incident field, but convergence is attained a few iterations sooner (e.g., compare iteration 4 of the first two columns). In the vicinity of π ($\pm\pi$, $\pm 1.1\pi$, and $\pm 1.2\pi$) the asymmetrical behavior is actually accentuated compared with that in the first column, though the reconstructions for using the exact internal field are decidedly more accurate. In fact, the reconstructions for phase shift $= \pm\pi$ are actually fairly acceptable as quantitative estimates of the speed of sound within the cylinder (especially the average). Comparison of iteration 1 of columns 1 and 2 shows a dramatic improvement in column 2 for all phase shifts having magnitude up to 1.2π. For $\Delta\psi = \pm 2\pi$ and $\pm 4\pi$, however, the reconstructions for all iterations in columns 1 and 2 are nearly identical, and settle upon erroneous estimated object functions.

Column three (rightmost) is unique in that the internal field is held fixed at its initial value, namely the exact internal field. In fact, higher iterations can be considered as various stages in one long iteration over Eqs. (3.1e.21). Very high quality reconstructions are obtained for phase shift up to $\pm 1.2\pi$. What is most remarkable, though, is the (very slow but definite)

convergence after 12 iterations for the case $\Delta\psi = \pm 2\pi$! Note that the first iteration reconstruction is identical to that in column 2 (because exactly the same calculation is performed). And that first iteration reconstruction is one with phase shift less than π (on the "wrong track"). It is currently not known why convergence to the correct solution was obtained for this case; it does not neatly fit the theory proposed in earlier sections. Further study could investigate this question; however, the question is merely academic, as in no practical setting will a guess for the field better than the incident field (let alone the exact field) likely be known. Furthermore, for $\Delta\psi = \pm 4\pi$ the solution in the third column is finally stable and erroneous, although of a different shape than that of the other two columns.

Finally, an aspect supportive of earlier statements that $\Delta\psi$ is fundamental in determining performance is clearly supported in all these plots. Especially notable is the strong comparablility of reconstruction quality between reconstructions of cylinders having phase shifts of equal magnitude and opposite sign. Also, it can be concluded from the failure of the algorithm for large $\Delta\psi$ when using the exact internal field that it is not the large error of the Born approximation that fundamentally limits performance of the sinc basis moment method, but rather the phase shift nonuniqueness problem described earlier in Section 5.7 and other factors. It also can be concluded that although high accuracy in the estimation of the internal field is indeed crucial for successful reconstruction (e.g., note the huge error for even small phase shift problems in iteration 1 of the first column, where

the internal field estimation is equal to the incident field), it is obtainable within a few iterations, at least for low magnitude phase shift problems (e.g., note the high accuracy of reconstructions in the 8th and 12th iteration of first column low contrast cases). Further study could investigate error in estimation of the internal field compared with the exact field for higher iterations of large contrast cases where the algorithm iterates on the field for which the incident field is the initial guess.

5.7f2 Initial guess for γ is close to γ^{ex}

In Fig. (5.7f2.1) is a composite showing iterations 9 and 12 (earlier iterations are similar) of a ka = 12.6, -27.2% speed of sound mismatch lossless cylinder ($\Delta\psi/\pi$ = 3.00) under the following conditions. In the left column the initial guess for the internal field was the incident field, and Eqs. (3.1e.22) were used in the iterations to improve that estimation, while in the right column the exact internal field was used as the internal field estimation for all iterations (no iterative corrections via Eqs. (3.1e.22) were made on the field). In the top row the initial guess for γ was zero; that is, the speed of sound was assumed to be c_0 everywhere within the object region. In lower rows, the initial guess for the speed of sound (i.e., through γ) was a cylinder of the exact shape of γ^{ex} but with the guess for the speed of sound within the cylinder, $\tilde{c}_1^0$, increasingly accurate for lower rows: 37.5% error (first row), 10%, 5.4%, 2.7%, and 0.0001% error. For the case of using the exact internal field, the algorithm settles on an erroneous solution for $\tilde{c}_1^0$ =

1.5 mm/μsec (= c_0), but gradually improves further down the composite to where the nearly perfect match at the bottom stays nearly perfect. Note that for the lower rows the solution only bends around $\tilde{c}_1^0$, and does not converge to c_1. There appear to be multiple weak local solutions highly dependent on $\tilde{c}_1^0$ for this high contrast case, which is unfortunate because it means that the initial guess must be very good to begin with, and then only minor improvement, if any at all, occurs from there. Regardless, as mentioned in the preceding subsection, knowledge of the exact internal field (let alone knowledge of the exact object shape) will not be available in a practical setting.

The situation is even worse in the left column, where the incident field is used as the initial guess for the field. All solutions are highly erroneous, even the one in the bottom row, for which the initial error in speed of sound in the cylinder was only 0.0001%. Perhaps it is this figure which is most disturbing and condemning for any future practical use of the sinc basis method, at least without some major basic improvements such as use of unwrapped phase in some way, or constraints as yet not applied to the iterating field and object function estimates. One additional further improvement that might be interesting to try would be use of the Rytov solution, as outlined in Chapter 2, as the initial guess for γ. However, as shown in this subsection, even a very good initial guess for the object function will immediately be corrupted if a poor estimate such as the incident field--the only one available--is used for the total internal field, so it probably would not be worth the effort. Furthermore,

as cited in Section 2.2c in the discussion following Eq. (2.2c.2), if $|f^{sc}|$ is of the order of or larger than $|f^{inc}|$, phase wrapping will cause problems in a practical implementation of the Rytov approximation. An initial study of magnitude of the scattered field vs. the incident field is presented in Table (5.7f2.1). In this example a cylindrical wave is scattered off a two wavelength radius cylinder of varying speed of sound contrast with respect to the coupling medium, while the scattered and incident fields are compared at two points on a line through the source and a diameter of the cylinder. Clearly, even for low contrast cases, the scattered field magnitude can be comparable or even greater than that of the incident field measured at the same location. The decline in percentage for a 100% contrast mismatch occurs because of a spreading out of the scattered field distribution for very high contrast cylinders (compare, for example, Figs. (5.5.1) and (5.5.3)). Thus, assumptions necessary for accurate reconstruction for moderate to high contrast objects with the Rytov solution are called into question from the outset.

5.7f3 Other initial guesses for γ

In this small subsection two types of initial guesses for γ that in a clinical situation could conceivably be available (in some inferior form) are investigated. First, knowledge of the speed of sound within the cylinder is assumed, but the boundary is uncertain. (In a practical situation, the speed of sound would also be uncertain.) Figure (5.7f3.1) illustrates the low contrast case: a -5% speed of sound mismatch, ka = 12.6 (m_{cyl} = 8)

cylinder. The initial guesses for γ are chosen to be (left to right and top to bottom) zero, and cylinders of radii 4, 6, 8, 10, and 12 pixels, within which the speed of sound is set to c_1. The initial guess, exact solution, and iterations 1 and 4 are shown. For initial guess cylinders both too small and too large, iteration 1 is the solution farther from γ^{ex}, and closer to $\tilde{\gamma}^0$ (except the cylinder of radius 8, which is $\tilde{\gamma}^0 = \gamma^{ex}$, and the incident field corrupts $\tilde{\gamma}^1$ away from $\tilde{\gamma}^0 = \gamma^{ex}$, and γ^4 again approaches γ^{ex}). The 12-pixel cylinder case is interesting in that the distortion outside the true 8-pixel radius is confined to oscillations about c_0, in contrast to the others, where the distortion extends about halfway from c_0 to c_1. Note that, contrary to the indications of Fig. (5.7g.1), the 12-pixel radius cylinder does not completely fill up the object region; it fills up only the midline cross sections, one of which is shown. Compared with guessing zero for γ, the other guesses yield comparable reconstructions by iteration 4, and are actually slightly inferior for the guess of the smallest cylinder.

Figure (5.7f3.2) shows the same study, for two fewer initial guess cylinder sizes, for intermediate and very high contrast speed of sound mismatches (c_1 = 1.35 mm/μsec and 1.263 mm/μsec, corresponding to $\Delta\psi/\pi$ = 0.89 and 1.5, respectively). The initial guesses for γ for both contrasts are, left to right, zero, and 6, 8, and 10 pixel radii cylinders with speed of sound set equal to c_1, so the third column from the left is the case $\tilde{\gamma}^0 = \gamma^{ex}$. The more symmetric, and closer to c_0 curves are iteration 1, and the more jagged, nonsymmetric curves are iteration 4. For c_1 = 1.35

mm/μsec, an accurate boundary definition in the reconstruction is evident, if not also accurate speed of sound estimation. But for the high contrast object, even beginning with $\tilde{\gamma}^0 = \gamma^{ex}$, the algorithm fails because $|\Delta\psi| > \pi$. Again, there is no substantial improvement over guessing zero for γ, after four iterations. In conclusion, while for low contrast cases (such as that in Fig. (5.7f3.1)) the sinc basis method corrects well for errors in structural estimations, no amount of knowledge of structure and/or speed of sound values can improve convergence conditions when the field is equal to the incident field, as noted in Section 5.7f2.

The other type of initial guess is an improved speed of sound initial estimation over guessing c_0, but no knowledge of structure ($\tilde{\gamma}^0$ is spatially uniform). Figure (5.7f3.3) shows speed of sound reconstructions of ka = 12.6, c_1 = 1.35 mm/μsec (upper row) and c_1 = 1.263 mm/μsec (lower row). The initial guesses are shown (horizontal lines indicating the level of uniform speed of sound initial estimation), the exact solution, and the first and fourth iterations. Again, the more symmetric and closer to c_0 curves are iteration 1, and the more jagged, nonsymmetric curves are iteration 4. On the left reconstructions, the speed of sound estimation for all pixels is c_0, and on the right reconstructions, $(c_0 + c_1)/2$. The results are indistinguishable, indicating the independence of $\tilde{\gamma}$ on the uniform initial estimate of speed of sound. Therefore, this is not a helpful technique for improving convergence quality over the case $\tilde{\gamma}^0 = 0$. A few additional alternative initial guesses for γ are considered in

the tables discussed in Section 5.14, but the following conclusion is unavoidable: for $\Delta\psi/\pi > 1$, failure occurs independent of the initial guess for the object function; only improvements in the internal field estimation help, and then only for small phase shift problems.

5.8 Lossy Cylinder Reconstructions

With the availability of a complex bessel function routine, COMBES, from Math Sciences Library, it was possible to modify the scattered field generator program to calculate scattered fields from lossy cylinders. Depending on the magnitude of the absorption coefficient, loss is in a sense minimizing multiple scattering effects because scattering that was before weak in the object region in a particular direction is with loss even further reduced in magnitude. Computationally, in the reconstructions this effect is manifested in a complicated way, through the coefficients, incident field, and modified scattered field (compared with the corresponding lossless case), which determine a modified estimated γ: first as estimated through the scattered field equations and then indirectly through a modified estimate of the field obtained from the internal field equations.

First, in Fig. (5.8.1) the speed of sound was varied for a ka = 12.6 cylinder while the pressure absorption coefficient was held constant at the very high value of 0.2 Np/mm at a frequency of 2MHz. The speed of sound percent contrasts are -5, +5, 6.7, 10., 13.3. As the speed of sound increased, the real part of the exact object function ranged from 2.1 to -5.3 times its imaginary

part. In spite of this, the speed of sound and absorption reconstructions both begin to be slowly converging and unacceptable at about the same level of speed of sound contrast. Note that for c_1 = 1.7 mm/μsec, while the reconstructed speed of sound steadily improves with the iterations, the reconstructed absorption for iteration 4 is much worse than that for iteration 1. However, by iteration eight (Fig. (5.8.2)) an improvement, though far from acceptable, is evident. This again indicates the importance of the accuracy of the estimation of the internal field for both speed of sound and absorption reconstructions, especially for high speed of sound and absorption mismatches.

In Figs. (5.8.3) and (5.8.4) are shown plots representing the reverse of 5.8.1 and 5.8.2: the speed of sound contrast is held constant at +5% while the pressure absorption coefficient varies as follows: 0, 0.1 Np/mm, 0.4 Np/mm, 0.7 Np/mm, 1.0Np/mm. Here a different behavior is observed: when the speed of sound reconstruction fails, the absorption reconstructions are still quite good after four (Fig. (5.8.3)) and especially after eight iterations (Fig. (5.8.4)). Several factors are influencing this behavior. The norm of the object function is increasing and consequently, so is the contribution of the object function to the widening of the object function-total field spectrum, whose width determines the estimated Nyquist sampling density. Also, the initial guesses for the object function and the internal field are increasingly worse. However, the phase shift problem described in Section 5.7 should not be present because absorption mismatches vary amplitudes, not phases, and the phase shift mag-

nitude through all these cylinders is far less than π--only 0.4π. Thus, although the maximum value of the imaginary part of the object function is greater than the maximum of the real part in Figs. (5.8.1) and (5.8.2), the shape and values of that absorption reconstruction are significantly better than those of the corresponding speed of sound reconstruction (comparing bottom plots on this and the previous composite). Also, the value of the imaginary part of γ is twice that of the real part for the bottom row reconstruction in Figs. (5.8.3) and (5.8.4), yet the absorption reconstruction is slowly closing in on the desired solution while the speed of sound reconstruction has settled on an erroneous solution. Therefore, speed of sound reconstructions may be more fundamentally limited than low speed of sound-mismatch absorption reconstructions, although in reality such factors as signal-to-noise ratio and others may dominate.

In Figs. (5.8.5) and (5.8.6) the speed of sound was set to c_0, a perfect match with that of the coupling medium, to examine the behavior of absorption reconstructions without including the phase shift effects. One finds results similar to when not superior to those of Figs. (5.8.3) and (5.8.4), which had the same absorption values but a 5% speed of sound mismatch. Again, for very high absorption values the reconstruction process becomes slow and accuracy steadily diminishes, but these absorption values are many times those that would be found in soft tissue, even at substantially higher frequencies within the diagnostic range. And under conditions of a perfect speed of sound mismatch, perhaps object size would not be such a limiting factor, at least

not theoretically by the phase shift requirement. (The fact that for the 8th iteration (Fig. (5.8.6)) the speed of sound reconstruction appears worse (larger values) than in the first iteration for the case α = 0.4 Np/mm is due to an artifact of the scaling of the plotting routine.) These reconstructions can be compared with speed of sound contrast reconstructions, where there the absorption was held constant at zero.

Finally, Figs. (5.8.7) and (5.8.8) show the reverse of Figs. (5.8.5) and (5.8.6): the absorption is a perfect match with the coupling medium, zero, while the speed of sound mismatch is again varied as -5%, 5%, 6.7%, 10%, and 13.3%. Here one finds that, for the same magnitude in object function, the speed of sound reconstructed cylinder shape is far worse than for that of the maximum contrast absorption reconstruction in Figs. (5.8.5) and (5.8.6). Also note an erroneous cylinder shape in the first iteration of the absorption reconstruction, which should be zero, whereas in Figs. (5.8.5) and (5.8.6) the Born approximation yielded speed of sound profiles which oscillated about the exact solution.

A general conclusion from this study might be that in the case of weak scattering, performance of speed of sound and absorption reconstructions are comparable, but because absorption contrast does not contribute to phase shifts while speed of sound contrast does, the algorithm will converge for stronger contrast and presumably larger size absorption objects than speed of sound mismatch objects. Of course in reality, <u>both</u> speed of sound and absorption contrasts will be significant. Note that for a lossy (0.02 Np/mm), very high speed of sound contrast cylinder ($\Delta\psi/\pi$ =

3.0), both speed of sound and absorption reconstructions failed. Thus, the phase shift problem destroys both speed of sound and absorption reconstruction quality.

5.9 Other Alternative Objects

5.9a Shifted cylinder

To demonstrate that the sinc basis moment method works for objects that are not circularly symmetric, a program was made to calculate exact scattered field data for a cylinder not centered with respect to the transducers and therefore not with respect to the object region origin, and the resulting data were given as input to the tomography program. The cylinder reconstructed was a ka = 14, 5% speed of sound mismatch cylinder. Reconstructions were compared for a cylinder centered with respect to the origin with that for a cylinder offset 5x5 pixels on the 25x25 grid. The result is shown in Fig. (5.9a.1) where speed of sound (above) and absorption reconstructions (below) are shown after four iterations for a centered cylinder (left) and an offset cylinder (right). The results are clearly for all purposes identical.

5.9b Contrived object

For the sake of investigating reconstruction quality for a more complicated object, a superposition of three simple lossy shapes was chosen. The values for c_1 and α_1 and sizes were chosen as follows:

shape	c_1(mm/μsec)	α_1(Np/mm)	size(wavelengths)
smaller cylinder	1.485	0.16	0.75 diameter
larger cylinder	1.515	0.32	1.50 diameter
rectangle	1.575	0.23	3.0x1.5

The largest phase shift, that through the diagonal of the rectangle, was consequently -0.3π, well within the range for successful reconstruction of cylinders. Of course, no exact solution was available for this multiple scattering problem, so the internal field equations were used to solve for the internal field, and then it and the exact γ were used in the measured scattered field equations to generate the scattered field. The results are shown in Fig. (5.9b.1); the exact γ in the upper row and the fourth iteration reconstruction in the bottom row. On the left are speed of sound distributions, and on the right, absorption. The reconstruction is quite close to the exact solution in appearance, and is also numerically accurate, as the center-line profiles in Fig. (5.9b.2) indicate (speed of sound in Fig. (5.9b.2a) and absorption in Fig. (5.9b.2b)), where in addition the first iteration has been included. Note that the behavior of the first iteration is similar to that for the circular cylinder reconstructions; the reconstructed speed of sound is too close to c_0, while the absorption reconstruction has a shape erroneously mimicking that of the exact speed of sound distribution. The fourth iteration, however, oscillates about the exact solution. Raising the phase shift through the diagonal of the rectangle above π predictably failed, as would be inferred from the discussions in Section 5.7.

5.10 Comparison of First Iteration of Sinc Basis Method with First-order Diffraction Tomography Solution

In this section attention is turned to the comparison of the sinc basis moment method with a representative first-order diffraction tomography algorithm based on the Born approximation and implemented by several researchers. This method was discussed in detail in Section 2.2. A short summary with illustrative figures follows. In the first-order Born solution, Fourier inversion method, based on the Fourier Diffraction Theorem, the Green function is expanded over an angular spectrum. That expansion is substituted into the integral equation for the scattered field. The coefficients of the resulting angular spectrum representation of the scattered field are directly expressable in terms of the Fourier transform of the field evaluated on a convenient surface outside the object region. If now the incident plane wave field is substituted as an approximation of the total field under this integral and the result is identified as the complex scattered field, the Born approximation has been made. And with this substitution a second identification of the angular spectrum coefficients can be made: they are values on a semicircular arc of the Fourier transform of the object function. (Had the substitution of the incident field for the total field not been made, these coefficients would represent Fourier transform values of the object function times the total, unknown, field, and no intrinsic property of the medium could then be ascertained.) The author implemented on the lab computer a diffraction tomography simulation program based on the above formulation using the first-order

Born approximation (as opposed to the Rytov approximation, which can also be used in the above theory to obtain a diffraction tomography algorithm). Signal processing details of (Pan and Kak, 1983) were followed closely. For example, coordinate transformations are needed for points on a semicircular arc from rectangular to polar coordinates, for retrieval of the closest available sample (specified on a polar grid). One of the two transformations (one for each half of the semicircular arc) is given here in the notation of Section 2.2. From Fig. (5.10.1) the following relations are easy to see:

$$k_u = k_0 \sin\beta$$

$$k = \sqrt{k_x^2 + k_y^2} = k_0 \sin\frac{\beta}{2} \rightarrow \beta = 2\sin^{-1}\frac{k}{2k_0}$$

or

$$k_u = k_0 \sin\left[2\sin^{-1}\left(\frac{\sqrt{k_x^2 + k_y^2}}{2k_0}\right)\right]. \qquad (5.10.1)$$

(Pan and Kak (1983) left off the leading k_0 in two equations for k_u.) And for the angle,

$$\phi - \frac{\pi}{2} = \theta + \frac{\beta}{2}$$

or

$$\phi = \tan^{-1}\left(\frac{k_y}{k_x}\right) + \sin^{-1}\left(\frac{\sqrt{k_x^2 + k_y^2}}{2k_0}\right) + \frac{\pi}{2}. \qquad (5.10.2)$$

Also, for use of the continuous space Eq. (2.2c.16) of Section 2.2c for reconstructions using discrete Fourier transforms, a factor of $h^2/\Delta u$ on the right-hand side must be included to account for the respective one- and two-dimensional discrete

integrations of the transforms, where h is the grid spacing and Δu is the receiver spacing. For a quick demonstration of the operation, the following problem is considered briefly. A ka = 18.8, 5% speed of sound mismatch ($\Delta\psi/\pi = -0.57$) cylinder is reconstructed, using 32 transmitters, a grid size of 64x64 with $\lambda/2$ spacing, and 512 receivers at spacing $\lambda/4$. Upon Fourier transformation of the scattered field and scaling with the phase factor found from Eq. (2.2c.16), the Fourier transform of γ is available on the loci shown in Fig. (5.10.2a). The real part of the numerical distribution of these values of γ (in folded form for FFT convenience) is shown in Fig. (5.10.2b). The real part of the Fourier transform of γ is shown in Fig. (5.10.3a), while the estimate of γ by this method (nearest neighbor interpolation of available values used) is depicted in Fig. (5.10.3b). This comparison makes clear how this approximation is a modified low-pass filtered version of the true object function. Clearly, the modifications are far more significant than mere low-pass filtering; the shape of $\tilde{\gamma}^{est}$ is significantly different from $\tilde{\gamma}^{ex}$ even for moderate spatial frequencies (e.g., second lobe). Finally, γ^{ex} (purely real) and the estimated γ are shown in Figs. (5.10.4a) and 5.10.4b, respectively.

It is obvious that this Fourier domain interpolation and inversion technique has very little in common computationally with the sinc-basis moment method, where closed-form integrations of quadruple sinc functions, matrices composed of elements having Hankel function sample values and numerical integrations of Bessel functions times double sinc functions as factors, and under-

relaxed iterative orthogonal projections using these matrices combine into a complex procedure alien to the Fourier Diffraction Theorem. Furthermore, the measurement geometry was quite different: the Fourier Diffraction Theorem in its commonly stated form requires the receivers to be on a line in front of (reflection mode) or behind (transmission mode) the scattering object, and they must be spaced closely enough to satisfy a Nyquist criterion. For the sinc basis moment method, receivers can be placed on any surface outside the object region, and there are no spacing requirements on these measurements, for no discrete Fourier transforms are taken of the received signals. Figure (5.10.5) shows center line profiles of speed of sound and absorption reconstructions of ka = 12.6 lossless cylinders for three contrast values: -10%, 5%, and 13.3% corresponding to $\Delta\psi/\pi$ = 0.9, -0.4, -0.94. The first iteration of the sinc basis moment method and the Fourier Diffraction Theorem are in all cases the curves very close together, the other two curves being the fourth iteration of the sinc basis method and the exact solution (straight lines). For this comparison study the receivers were placed on a line perpendicular to the direction of the incident wave and behind the object region in the Fourier Diffraction Theorem algorithm, and left as before on a ring around the object region in the sinc basis algorithm. The only things kept in common between uses of the two methods were the object region grid spacings, the distance from the receiver line to the center of the object region in the Fourier domain method and the radius of the ring of detectors in the sinc basis method, the incident fields, the

object functions being reconstructed, and the use of the Born approximation in the integral equation for the scattered field. The common use of the Born approximation is the cause for the astounding agreement between the two solutions for both low and high magnitude, positive and negative contrast cases. What is most amazing is that they both fail in exactly the same way for the high contrast cases where it is known that the Born approximation is not valid.

Table (5.10.1) provides further proof of similarity between the two solutions for the three cases examined above. Define $\gamma^{S.B.,1}$ to be the estimated object function after one iteration of the sinc basis moment method and $\gamma^{F.D.T.}$ to be the estimated object function resulting from the Born approximation mode of the Fourier Diffraction Theorem method. The complex normalized inner products between $\gamma^{S.B.,1}$ and γ^{ex}, $\gamma^{F.D.T.}$ and γ^{ex}, and $\gamma^{S.B.,1}$ and $\gamma^{F.D.T.}$ are given for the three contrast values. For the high contrast cases $\gamma^{S.B.,1}$ and $\gamma^{F.D.T.}$ are both far from one but nearly equal; moreover $\gamma^{S.B.,1}.\gamma^{F.D.T.}$ is practically equal to one for all cases. This connection allows assessment of the sinc basis moment method and in particular gives strong credibility to a generalization of the applicability of the conclusions made earlier using the error plots discussing first vs. higher iterations of the sinc basis method to the comparison of the Born approximation vs. higher iterations of the sinc basis method.

5.11 Use of Plane Rather than Cylindrical Wave for Incident Field

Any incident field satisfying the homogeneous Helmholtz wave equation is valid to use for f^{inc} in the reconstruction equations. To demonstrate this fact numerically, a plane wave was used for the incident field, defined as follows. If ϕ_t and ϕ_j are the angles the transmitter and pixel j, respectively -to-center of object region lines make with the x-axis, then the propagation distance to pixel j, referenced to the origin, is

$$r_0 = -r_j \cos(\phi_t - \phi_j), \tag{5.11.1}$$

the minus sign due to an additive π within the cosine. Thus, the plane wave incident field is

$$f_{\phi j}^{inc} = e^{-jk_0 r_0} \tag{5.11.2}$$

The results for a ka = 0.75, 5% mismatch cylinder were essentially identical to those using a cylindrical wave. For example, the squared errors in $\tilde{\gamma}$ compared with those using an incident cylindrical wave are

	$\sum_j (\tilde{\gamma}_j{}^i - \gamma_j{}^{ex})^2$	
iter. i	plane wave	cylindrical wave
0	2086.2	2086.2
1	305.7	307.0
2	155.8	155.8
3	141.9	141.3
4	140.5	139.8

(Note that the series solution for the scattered field for an incident plane wave scattering off a cylinder differs from that used here for an incident cylindrical wave; the difference in the series coefficients is discussed in Section 3.2d1.)

5.12 Use of Line Rather than Ring Geometry for Receivers

It was mentioned in Section 3.3c that any convex surface can in theory be used for measurement of the scattered field which serves as input data to the sinc basis moment method. Figure (5.12.1) shows an alternative geometry to that used in Chapter 3 and shown in Fig. (3.1e.2); here the receivers, rather than being placed on a ring of radius R from the center of the object region, are placed on a line opposite the transmitting transducer and out a distance R from the object region center. For scattered field generation two parameters are necessary: the receiver-to-object (cylinder) center distance, which from Fig. (5.12.1) is $\sqrt{R^2 + u^2}$, and the angle between the transmitter-to-object region center line and the object region center-to-receiver line, which from Fig. (5.12.1) is $\pi - \tan^{-1}(u/R)$. For the coefficients of the scattered field equations, D_{mj}, all that is needed is the distance from the pixel under consideration to the receiver, r, which from Fig. (5.12.1) is

$$r = \sqrt{R^2 + u^2 + r_j^2 - 2\sqrt{R^2 + u^2}\, r_j \cos\phi} \qquad (5.12.1)$$

where

$$\phi = \pm[\phi_j - \phi_t + \pi + \tan^{-1}(\frac{u}{R})] \qquad (5.12.2)$$

where ϕ_j and ϕ_t are the respective angles of the pixel and transmitter coordinate vectors measured with respect to the x axis, r_j is the pixel-to-object region center distance, and u is the value of the u coordinate for the receiver along the measurement line. The choice of plus or minus depends upon the relation between receiver, pixel, and transmitter positions, but in either case,

$$\cos\phi = -\cos[\phi_t - \phi_j - \tan^{-1}(\frac{u}{R})]. \qquad (5.12.3)$$

The receiver spacing Δu was chosen to match the spacing of detectors for the same value of R for the detectors on a ring: $2\pi R/n_{det}$. All programs (tomography and one of its subroutines, coefficient generation, and scattered field generation) had to be modified. For the D_{mj}, no attempt was made to exploit symmetry that might greatly reduce the number of coefficients needed. The situation in this regard is even more serious than coefficient storage for the detectors on a ring, because for each view the receiver coordinates change, and therefore, so do the distance relations between receivers and pixels. Thus, D_{mj} now also depend on the viewing angle, and so must be calculated and stored separately for each view; here we need three subscripts for the coefficients: $D_{mj\phi}$. Exploiting symmetry could probably reduce the number of $D_{mj\phi}$ to an acceptable level, but this project was not considered worth the effort in this study; consequently, on the

VAX the grid size for this case was limited to 11x11. For the 11x11 case, comparable results were obtained compared with the corresponding reconstruction carried out with receivers on a ring. For a ka = 0.75, 5% speed of sound mismatch, receiver spacing for both cases of 3λ, the squared error in $\tilde{\gamma}$ was

$$\frac{\sum_j (\tilde{\gamma}_j^i - \gamma_j^{ex})^2}{}$$

iter. i	ring	line
0	2086.2	2086.2
1	307.0	325.0
2	155.8	212.0
3	141.3	206.0
4	139.8	204.0

(Note that for this particular case of object circular symmetry, the received field on a line is the same for all views; in the ring geometry, the same data are merely permuted from view to view.)

5.13 Execution Times

Table (5.13.1) summarizes execution times for four and eight iterations of the sinc basis moment method performed on the Alliant supercomputer. Those values are a far cry from the initial reconstructions on the VAX where an 11x11 reconstruction took 17 hours! The execution times after eight iterations are plotted in Fig. (5.13.1) against a best-fitting an^4_{max} curve, where a is a constant ($2.03 \cdot 10^{-6}$) vs. n_{max}. The strong matching

of the n^4_{max} dependence, as predicted in Chapter 4, is quite clear.

5.14 Additional Numerical Comparison Studies

Further knowledge about the characteristics and behavior of the sinc basis moment method can be obtained by controlled experiments within the computer simulation context. In Chapter 3, such parameters as frequency, object contrast, object size, sampling density, iteration duration, grid size, relaxation parameters, and order and specific form of ART-type corrections were idependently varied, and a basic outline of the behavior of the algorithm for small objects resulted. In Chapter 4, the method of carrying out the internal field equations convolutions was altered, and identical results were obtained with a substantial reduction in computational complexity. In earlier sections of this chapter, variation in method of solution of the scattered field equations was qualitatively discussed (the alternative methods being QR decomposition (Section 5.3a) and the ART-type algorithm discussed in Section 5.3b), the form of scattered field data (exact series solution vs. sinc basis equations in forward mode), order of corrections in the scattered field equations, object size for larger objects, object contrast for larger objects, phase shift through the object, type of initial guess for the internal field, absorption coefficient within the cylinder, method of solution of the inverse scattering problem (sinc basis method vs. the Born approximation mode of the Fourier Diffraction Theorem method), type of incident field (plane vs.

cylindrical waves), and receiver geometry (line vs. ring) were all investigated either quantitatively or qualitatively. In this final section, a few of the same parameters which were earlier discussed qualitatively will be approached more quantitatively through the use of tables, and a few more variations will be investigated. The contents of the tables to be described in this section are far too numerous to be given a comprehensive and uniform analysis. For example, the specific sets of conditions for which a particular reader would like to compare results may vary tremendously; only a few will be pointed out here. Mainly, this discussion will identify the columns and rows sufficiently clearly so that the reader may easily look up desired information contained within the table, although some analysis is given.

5.14a Contrast variation for ka = 4.7, 12.6, 32, and 46 cylinders

Tables (5.14a.1), (5.14a.2), (5.14a.3), and (5.14a.4) comprise data from the object contrast runs for which several plots of various errors and improvement factors were presented in Section 5.7d. The only parameter changing between these tables was the object size: for Tables (5.14a.1), (5.14a.2), (5.14a.3), and (5.14a.4) the object sizes are ka=4.7, 12.6, 32, and 46 (m_{cyl}=4, 8, 18, and 25). The columns from left to right in Tables (5.14a.1a), (5.14a.2a), (5.14a.3a), and (5.14a.4a) are as follows: the speed of sound within the cylinder c_1; percent contrast; phase shift through cylinder $\Delta\psi/\pi$; squared error in estimated γ for iteration i $E_i = \sum_j (\tilde{\gamma}_j{}^i - \gamma_j{}^{ex})^2$ for i = 0, 1, 4,

8; log of normalized squared error in $\tilde{\gamma}$ for iteration i $S_i = \log_{10}\{E_i/n_{\gamma\neq 0}\}$ for i = 0, 1, 4, 8; improvement factor in $\tilde{\gamma}$ from iteration i to iteration j $S_{ij} = 10\log_{10}\{E_i/E_j\}$ for ij = 01, 04, 08, 14, 18. In Tables (5.14a.1b) through (5.14a.4b) are listed, left to right, the speed of sound within the cylinder c_1; the average of the estimated speed of sound within the cylinder region for iteration i $\tilde{\tilde{c}}_1{}^i$ for i = 1, 4, 8; the percent error for that average estimate for iteration i $e_1{}^i$ for i = 0, 1, 4, 8 (for iteration zero, this is $100(c_0 - c_1)/c_1$, as opposed to the percent contrast column in Tables (5.14a.1a) through (5.14a.4a), which is $100(c_1 - c_0)/c_0$); the average of the estimated speed of sound outside the cylinder but within the object region for iteration i $\tilde{\tilde{c}}_0{}^i$ for i = 1, 4, 8; the percent error for that average estimate for iteration i $e_0{}^i$ for i = 1, 4, 8 (for iteration zero, this error is zero). In Tables (5.14a.1c) through (5.14a.4c) are listed, left to right, the speed of sound within the cylinder c_1; the average of the estimated absorption coefficient within the cylinder for iteration i $\tilde{\tilde{\alpha}}_1{}^i$ for i = 1, 4, 8 (all cylinders are lossless); the average of the estimated absorption coefficient outside the cylinder (but inside the object region) for iteration i $\tilde{\tilde{\alpha}}_0{}^i$ for i = 1, 4, 8 (the coupling medium is lossless).

For all object sizes the contrasts were chosen such that the phase shifts $\Delta\psi$ varied from approximately 2π to -2π. Comparing the early with the late errors either via the E_i or the S_i, one sees a dramatic change of behavior between approximately $|\Delta\psi| = 0.8\pi$ and $|\Delta\psi| = \pi$. The comments appearing in Section 5.7d in reference to the graphs plotting these results can be verified

here numerically. Note that the threshold, $|\Delta\psi| = \pi$, corresponds to different percent contrasts for the different sized cylinders: as high as -25%/50% for the ka = 4.7 cylinder (Table (5.14a.1a)) down to only -3.3%/3.5% for the largest cylinder, ka = 46 (Table (5.14a.4a)). This again points toward using $\Delta\psi$ rather than c_1 as the threshold parameter for reconstruction quality. Also plotted and discussed in Section 5.7d were the S_{ij}, which for a given object size can all be more conveniently compared at a glance in Tables (5.14a.1a) through (5.14a.4a). For example, one sees for the ka = 32 cylinder (Table (5.14a.3a)), at $\Delta\psi/\pi = 0.7$, the striking increase in improvement factor over the initial error from $S_{01} = 0.9$ to $S_{04} = 13.5$ to $S_{08} = 16.9$, and over the Born approximation from $S_{14} = 12.6$ to $S_{18} = 16.0$. Incidentally, $S_{ij} = S_{1j} - S_{1i}$, where here 1 = 0, i = 1, and j = 4 or 8, because of the property of the log of a quotient.

For direct comparison with c_1, the average estimate of c_1 within the cylinder is given in columns 2, 3, and 4 of Tables (5.14a.1b) through (5.14a.4b). The unchanging pattern for all sizes is that initially (after one iteration--the Born approximation) the average reconstructed speed of sound is too close to c_0, as was seen in the center-line profiles presented in earlier sections. The threshold at $\pm\pi$ is again dramatic. After eight iterations, for the ka = 46 cylinder, percent errors in $\tilde{\bar{c}}_1$ are less than 0.07% for all $\Delta\psi$ less than π (to about $\pm 0.7\pi$); but for $\Delta\psi = \pm\pi$, the percent errors jump to more than 1%. Somewhat easier for at-a-glance comparisons are the percent errors of these estimates. Again, the tendency is noticed for the recon-

structed speed of sound to be too close to c_0: in iteration one the errors are opposite in sign to the percent contrasts (listed in Tables (5.14a.1a) through (5.14a.4a)). Also, the sharp thresholds displayed in the graphs of Figs. (5.7d.11), (5.7d.12), (5.7d.15), and (5.7d.16) are obvious in these columns. Not shown in any graphs were the average estimated speed of sound outside the cylinder but within the object region, and the corresponding percent errors, all given in the remaining columns of Tables (5.14a.1b) through (5.14a.4b). Because the initial guess was perfect (c_0), these errors all tend to be quite low (in comparison with those for c_1). What is noticeable is that except for the smallest cylinder (ka = 4.7, Table (5.14a.1b)), where the errors in $\tilde{\tilde{c}}_1{}^i$ become as large as 20.4% for the -40.0% contrast case, the errors in $\tilde{\tilde{c}}_0{}^i$ even for the highest contrast cases are almost all of magnitude less than one percent. For example, for ka = 12.6 (Table (5.14a.2b)), -15.8% contrast, the error in the average reconstructed speed of sound within the cylinder rises from 23.3% to 27.4% from iteration 1 to 8, while the error in the reconstructed speed of sound outside the cylinder decreases (in magnitude) from -0.33% to -0.04%. The fact that the outside border matches with the unbounded homogeneous space, in addition to the perfect initial guess, causes the maintenance of quality of reconstructed speed of sound to be quite robust in that region. Except for a very few of the most extremely high contrast cases, the error at iteration 8 is significantly lower than that of iteration 1.

Also decreasing over the iterations, except for a few very high contrast cases (mostly for the smallest cylinder), is the reconstructed absorption coefficient within the cylinder (which in all cases should be zero). But the same phenomenon as happened with the speed of sound happens here: for the high contrast cases, the reconstructed absorption coefficient within the cylinder becomes far worse with increasing speed of sound contrast, while the reconstructed absorption coefficient outside the cylinder region remains far lower even out to the highest contrast cases. It is of passing interest to note that absorption coefficients must always be positive, and that while $\tilde{\tilde{\alpha}}_1{}^i$ are nearly all positive, $\tilde{\tilde{\alpha}}_0{}^i$ are slightly negative roughly for $|\Delta\psi| < \pi$ and positive for $|\Delta\psi| > \pi$.

5.14b Sequence of controlled experiments for small cylinders

Tables (5.14b.1a) through (5.14b.1d) comprise a series of controlled experiments carried out by computer simulations. In these runs, $n_{max} = 25$, $m_{cyl} = 8$, ka = 12.6, $\beta_0 = 1.0$ (see Eq. (3.2g.1)), and zero noise was added to the scattered field data to eliminate one source of reconstruction error. Generally, from experiment to experiment one parameter only was varied. The following are varied in Tables (5.14b.1a) through (5.14b.1d): the scattered field generation (sinc basis method using γ^{ex} and $f^{int,ex}$ in Eqs. (3.2e.21)) vs. the exact series solution (see Eqs. (3.2d1.8) and (3.2d1.24) through (3.2d1.29)); speed of sound within the cylinder, c_1; the initial guess for γ (a variety was investigated); the initial guess for the internal field (the

incident field (Eq. 3.2a.1)) vs. the exact series solution (see Eqs. (3.2d1.9) and (3.2d1.25) through (3.2d1.30)); the relaxation parameter β_1 as defined in Eq. (3.2f.1); reordering vs. not reordering of the scattered field equations for minimum adjacent row inner products (see Section 5.4); and the overdetermination factor Q_2 as defined in Eq. (3.2g.2). The iteration numbers are indicated from zero (initial guess) to the final iteration number (between 4 and 12 inclusive). The following performance indicators were recorded: the squared error in $\tilde{\gamma}$ compared with γ^{ex} for iteration i, $E_i^{ex} = \sum_j (\tilde{\gamma}_j^i - \gamma_j^{ex})^2$; the additive improvement in the squared error in $\tilde{\gamma}$ compared with γ^{ex} from iteration zero to iteration one, $\Delta E^{ex} = E_1^{ex} - E_0^{ex}$ (a negative number indicates a decrease in squared error); the normalized inner product of $\tilde{\gamma}^i$ and γ^{ex}, $L_i^{ex} = (\tilde{\gamma}^i \cdot \gamma^{ex*}) / (|\tilde{\gamma}^i||\gamma^{ex}|)$; its magnitude $|L_i^{ex}|$; and, for the same runs, the same series of indicators, but all comparing $\tilde{\gamma}$ with the solution for γ obtained by solving the scattered field equations (Eqs. (3.1e.21)) by QR decomposition: $E_i^{QR} = \sum_j (\tilde{\gamma}_j^i - \gamma_j^{QR})^2$, $\Delta E^{QR} = E_1^{QR} - E_0^{QR}$, $L_i^{QR} = (\tilde{\gamma}^i \cdot \gamma^{QR*}) / (|\tilde{\gamma}^i||\gamma^{QR}|)$, and $|L_i^{QR}|$.

A description of the experiments begins with a statement of conditions for the first experiment: $f^{sc} = f^{sc,s.b.}$ (sinc basis equations), $c_1 = 1.575$ mm/μsec, $\tilde{\gamma}^0 = \gamma^{ex}$, $\tilde{f}^{int,0} = f^{int,ex}$, $\beta_1 = 0.2$, no reordering, and $Q_2 = 1.05$. Note that for $Q_2 = 1.05$ there are approximately twice as many equations as unknowns. Because the initial guesses for both γ and f^{int} are the exact solutions, there is zero error at iteration zero. The fact that $E_1^{ex} = 0$ indicates that γ^{ex} is at least a strong local solution

of the scattered field equations. In fact, it is the true solution, because the very same equations are used to generate the scattered field (left-hand side of Eqs. (3.1e.21)) as are used in the solution. The only reason the error does not stay zero indefinitely ($\tilde{\gamma}^i = \gamma^{ex}$ for all i) is that the internal field equations are <u>not</u> a perfect match; Eqs. (3.1e.22) were not used to generate the left-hand side (the incident field). So $E_1^{ex} = 7.6$ is an indication of the discretization error present in the internal field equations. But even after four iterations the squared error, 28.1, is only one three-hundredth of the squared norm of γ^{ex}. Naturally, the inner product is also perfect: $L_0^{ex} = L_1^{ex} = 1.00$. Now γ^{QR} must be more carefully defined. As mentioned above, γ^{QR} is the solution for γ of the scattered field equations by QR decomposition; it must be added here that there the scattered field was generated by the exact series solution given in Section 3.2d1, which is not a perfect match with the scattered field equations (Eqs. 3.1e.21), and that the internal field estimate was set equal to the exact internal field. Thus, $\gamma^{QR} \neq \gamma^{ex}$, and for this case that discrepancy is huge: E_0^{QR} for the first experiment is $\sum_j (\gamma_j^{ex} - \gamma_j^{QR})^2 = 8079841$, a thousand times $||\gamma^{ex}||^2$. Furthermore, $|L_0^{QR}|$, which for the first experiment is the magnitude of the normalized inner product of γ^{ex} and γ^{QR}, is only 0.03 (for colinear vectors this inner product would be 1.0, and for orthogonal vectors, 0.0).

In the second experiment the only change was that $\tilde{\gamma}^0$ was set to zero. The correction in four iterations towards γ^{ex} is marked, as all the indicators show; even L_i^{QR} approaches

$(\gamma^{ex} \cdot \gamma^{QR*}) / (|\gamma^{ex}||\gamma^{QR}|)$. Note that after only one iteration, by far the greatest improvement of those of all iterations was made in both E^{ex} and L^{ex}.

In experiment 3, Q_2 was reduced to 0.02 (approximately the same number of equations as unknowns). As one would expect, there is a significant decline in performance; E_i^{ex} are roughly three times those for the case $Q_2 = 1.05$. However, the error is still acceptable, and $|L^{ex}|$ becomes as high as 0.983.

Next, Q_2 was reset to 1.05 and the incident field was used as the initial guess for the internal field, $\tilde{f}^{int,0}$. Comparing with experiment 2, where the exact internal field was used, it is seen that ΔE^{ex} is much larger for using $f^{int,ex}$ than using f^{inc} (-7818 vs. -5660); but by iteration 4 the errors E_4^{ex} are nearly the same (99.1 vs. 122.9). One could conclude that for iteration 1 the large difference is a consequence of using the poor estimate for the internal field (f^{inc}) but by iteration 4 the internal field estimate has been sufficiently corrected for the estimates of γ in experiments 2 and 4 to be close. The same observation can be made about the L_i^{ex}.

Experiment 5 is number 4 repeated for the case $Q_2 = 0.02$. At first the errors for the two values of Q_2 are closer than are those of experiments 2 and 3, but by iteration 4, again the error for $Q_2 = 0.02$ is three times that for $Q_2 = 1.05$. Thus, overdetermination is of significant benefit even in the simulation environment.

For the remainder of experiments the scattered field was set to the Bessel function series solution described in Section 3.1d.

Thus, experiment 6 is directly comparable to number 1, where the exact solutions for both γ and f^{int} are used as the initial guesses. Because f^{sc} was no longer directly calculated by the sinc basis equations, there is no longer a perfect match between left- and right-hand sides of Eqs. (3.1e.21), so $E_1^{ex} = 37.4$ ($\neq 0$). Interestingly, $\tilde{\gamma}$ moves away from γ^{ex} and towards γ^{QR} ($\Delta E^{QR} = -130$), but after four iterations $\tilde{\gamma}$ is still very close to γ^{ex} and very far from γ^{QR}.

Next, the initial guess for γ was set to zero. The results are roughly comparable to those of experiment 2, except here the error is higher due to use of the exact scattered field. Also, it is interesting that ΔE^{ex} and ΔE^{QR} are quite close, yet $|L_1^{ex}|$ and $|L_1^{QR}|$ are nearly as far apart as possible for normalized inner product magnitudes. Thus, starting with $L_0^{ex} = L_0^{QR} = 0.0$, the subsequent $|L_i^{ex}|$ and $|L_i^{QR}|$ approach, respectively, one and zero ($\tilde{\gamma}$ moves toward γ^{ex}).

In experiment 8, Q_2 was set to 0.02, resulting in a predictably significant but acceptable degradation (E_i^{ex} still decreases from 8302 into the low hundreds by iteration 4).

In experiment 9, Q_2 was reset to 1.05 and $\tilde{f}^{int,0}$ was changed to f^{inc}. Nearly identical behavior as described for the case $f^{sc} = f^{sc,s.b.}$ (experiments 2 and 4) occurs here, though again the errors are shifted somewhat higher due to using $f^{sc} = f^{sc,ex}$. That is, ΔE^{ex} is far larger for $\tilde{f}^{int,0} = f^{int,ex}$ than for $\tilde{f}^{int,0} = f^{inc}$, but by four iterations the E_4^{ex} are quite close, demonstrating the ability of the sinc basis moment method to correct the internal field estimate.

Next, when Q_2 is set to 0.02, approximately the same ratio, 1.5, exists between the squared errors in $\tilde{\gamma}$ for the two values of Q_2 (comparing this experiment, number 10, with the previous experiment).

The scattered field equations relaxation parameter β_1 was varied in experiments 11 and 12 (and Q_2 was reset to 1.05). For $\beta_1 = 0.5$ (experiment 11) there is a slight improvement over the case $\beta_1 = 0.2$ (experiment 9) for this particular object size and contrast. It was stated in Section 3.3i that $\beta_1 = 0.2$ was, overall, a near-optimal value. That statement will be reverified in experiment 14, where using $\beta_1 = 0.5$ for a different problem leads to a highly deteriorated result, far outweighing the slight gain seen here for this low contrast case. In experiment 12, $\beta_1 = 1.0$; even for this low contrast case ($c_1 = 1.575$ mm/μsec) by iteration 4, there is an increase by a factor of 1.7 in E^{ex} over that for $\beta_1 = 0.2$, as well as a decline in $|L^{ex}|$. It is interesting to note that the behavior discussed with respect to γ^{ex} is paralleled in that with respect to γ^{QR}, especially for E^{QR} and ΔE^{QR} compared with that for E^{ex} and ΔE^{ex}.

In experiments 13 through 16, a larger contrast value, $c_1 = 1.65$ mm/μsec, was used to bring out certain behavior more clearly. For example, in experiment 13, $L_1{}^{ex}$ has both large real and imaginary parts, analogous to the mixing of real and imaginary parts of γ in $\tilde{\gamma}^1$ as discussed in earlier sections. By iteration 4, the real part of L^{ex} dominates, and a qualitatively good reconstruction is also obtained. Note that after only one iteration of correction to the field, the real part of L^{ex}

increases from 0.48 to 0.92, while its imaginary part decreases from 0.80 to 0.27. However, $|L^{ex}|$ increases only from 0.94 to 0.96, indicating the insufficiency of using $|L^{ex}|$ for assessing convergence quality. Of further interest is the fact that although ΔE^{ex} is only 14% of $E_0{}^{ex}$, $|L^{ex}|$ changes from zero all the way to 0.94 in one iteration. (In the corresponding low contrast case, experiment 9, ΔE^{ex} was nearly 67% of $E_0{}^{ex}$.) The indicator best reflecting qualitative reconstruction quality is an examination of both the real and imaginary parts of $L_i{}^{ex}$; one sees a substantial improvement over $\tilde{\gamma}^0 = 0$ in $L_1{}^{ex}$, yet one can also infer the erroneous mixing of real and imaginary parts of γ.

For a reverification of the near-optimality of the choice $\beta_1 = 0.2$ over a wide range of problems, β_1 was again set to 0.5 in experiment 14. The result was a squared error in $\tilde{\gamma}$ 2.7 times that for the case $\beta_1 = 0.2$; thus, considering a range of problems rather than an isolated case (e.g., experiment 11), the use of $\beta_1 = 0.2$ for most of the reconstructions presented in this thesis is justified.

In experiment 15, the value of alternating between correction of $\tilde{\gamma}$ and of $\tilde{f}^{int}$ is quantitatively addressed. Here, the scattered field equations were cycled through four times instead of once (the usual case). The use of four was chosen to reflect the number of times the scattered field equations are cycled through in a typical four-iteration run. Comparing $E_4{}^{ex}$ of experiment 13 with $E_1{}^{ex}$ of experiment 15 shows the difference between iteratively correcting the internal field estimate or not. The error $E_4{}^{ex}$ of experiment 13 is about one fourteenth of

E_1^{ex} of experiment 15. Clearly, iterating on the field for this moderate contrast case is crucial for obtaining an adequate reconstruction. One might argue that the only reason for the difference in performance is the increased amount of computation in the former case, but even if E_1^{ex} of experiment 13 is compared with E_1^{ex} of experiment 15 (where four times as much computation is represented in experiment 15), the conclusion is the same: the squared error in $\tilde{\gamma}$ is higher if more cycles through the scattered field equations are performed, without corrections on the field. All of the above analysis can be applied with equal validity to the L_i^{ex} for these two experiments.

Experiments 16 and 17 are respectively the high and low contrast cases of using the reordering of scattered field equations option, as described in Section 5.4. There is actually a small deterioration in both cases, despite the huge increase in computation time (E_4^{ex} = 1988 for reordering vs. 1938 (Experiment 13) for no reordering, for c_1 = 1.65 mm/μsec and 275 vs. 269 (Experiment 9) for no reordering for c_1 = 1.575 mm/μsec).

For the remainder of experiments except the last (18 through 25) the initial guess for f^{int} was $f^{int,ex}$, and the initial guess for γ was varied. For these experiments it was thought best to use the exact internal field as the initial guess, thereby partially eliminating one source of error that could obscure other behavior. (One might argue that $f^{int,ex}$ should be used for all iterations, but then observed behavior would not reflect normal use of this algorithm, where a major portion of the cycle is spent on and behavior determined by the correction

of the internal field estimate.) Experiments 6 and 18 through 22 investigate cases where $\tilde{\gamma}^0$ is set to a vector some fraction of the way between γ^{ex} and γ^{QR}, inclusive of γ^{ex} and γ^{QR}. Because of interest in the continued movement of $\tilde{\gamma}$ with the iterations, more iterations were allowed and recorded for these experiments: up to 12. In experiment 6 (discussed earlier), $\tilde{\gamma}^0$ was set to γ^{ex}. There, because of the discretization error in the scattered field equations, the solution estimation moved slightly away from γ^{ex}, in the direction of γ^{QR} (which is reasonable, because γ^{QR} is in fact the true solution of the "mismatched" scattered field equations, using the exact scattered field).

The average of γ^{ex} and γ^{QR} was used for $\tilde{\gamma}^0$ in experiment 18. Here, significant reductions in E^{ex} occurred up to about iteration 9, the solution thereafter relatively stabilizing. For this case, by definition the squared errors $E_0{}^{ex}$ and $E_0{}^{QR}$ were equal. Although the initial movement was slightly away from γ^{ex} and slightly toward γ^{QR}, subsequent corrections on the field worked in favor of the reduction of E^{ex} and increase in E^{QR}. However, $|L^{ex}|$ never approached one, and the reconstructions qualitatively were very poor. Note that $L_0{}^{QR} \cong 1$, so in orientation, $\tilde{\gamma}^0 = (\gamma^{ex} + \gamma^{QR})/2$ is very nearly in the same hyperdirection as γ^{QR}. This is because the squared norm of γ^{QR} is nearly a thousand times that of γ^{ex}. Thus, to put $\tilde{\gamma}^0$ midway between γ^{ex} and γ^{QR} in a directional sense, $\tilde{\gamma}^0$ should be taken as

$$\tilde{\gamma}^0 = \frac{1}{2}\left(\frac{||\gamma^{QR}||}{||\gamma^{ex}||}\gamma^{ex} + \gamma^{QR}\right)$$

$$= \frac{1}{2}\left(31.2\gamma^{ex} + \gamma^{QR}\right). \qquad (5.14b.1)$$

The results of using this initial guess for γ are shown in experiment 19. In this case $L_0^{ex} \simeq L_0^{QR}$, as opposed to experiment 18 where $|L_0^{ex}| = 0.063$ and $|L_0^{QR}| = 0.9995$. Both E_0^{ex} and E_0^{QR} are greatly increased over those for $\tilde{\gamma}^0 = (\gamma^{ex} + \gamma^{QR})/2$, but so are ΔE^{ex} and ΔE^{QR} (in fact these respective improvements toward γ^{ex} and γ^{QR} are essentially equal), so that E_1^{ex} and E_1^{QR} are comparable to their values in experiment 18. After correction of the field (in iteration 1), the behavior is much the same as that in experiment 18, indicating the dominating influence of internal field corrections on long-term behavior.

An interesting case is shown in experiment 20, where the initial guess for γ is directionwise oriented 90% towards γ^{ex} and 10% towards γ^{QR}: $\tilde{\gamma}^0 = 0.91(31.2\gamma^{ex}) + 0.1\gamma^{QR}$. Thus, the inner product magnitude $|L_0^{ex}| = 0.994$, while $|L_0^{QR}| = 0.14$. Note how the ART correction reduces the squared error in $\tilde{\gamma}$ but does not maximize the inner product $|L_i^{ex}|$. The simultaneous, nearly equal reductions in E^{ex} and E^{QR} actually, in a directional sense, move $\tilde{\gamma}$ toward γ^{QR} and away from γ^{ex}. Further iterations continue to reduce E^{ex}, but increase E^{QR}, while weakly maintaining the earlier trends in L_i^{ex} and L_i^{QR} in experiment 19. From this and earlier examples, it is clear that neither E_i nor L_i tells the

complete story of iteration quality, but in conjunction give a variety of indications.

In experiment 21, $\tilde{\gamma}^0$ was directionwise placed 90% towards γ^{QR} and 10% towards γ^{ex}; thus $|L_0{}^{ex}| = 0.14$ and $|L_0{}^{QR}| = 0.994$. Again, $\Delta E^{ex} \simeq \Delta E^{QR}$ and again $|L_1{}^{ex}| < |L_0{}^{ex}|$, while $|L_1{}^{QR}| > |L_0{}^{QR}|$. After the first correction of the field, there is a strong move evidenced both in E and L away from γ^{QR} and towards γ^{ex}, but thereafter not much activity except the incongruous simultaneous decrease in both $E_i{}^{ex}$ and $|L_i{}^{ex}|$ (again, evidence that ART minimizes squared errors and does not maximize inner products).

Finally, in experiment 22, $\tilde{\gamma}^0 = \gamma^{QR}$. Using ART to correct $\tilde{\gamma}$ produces essentially zero change in $\tilde{\gamma}$ (and $L_0{}^{QR} = L_1{}^{QR} = 1.00$), indicating that γ^{QR} is a strong (and known to be the true) solution of the scattered field equations for $f^{sc} = f^{sc,ex}$. Only after the internal field estimate has been modified does $\tilde{\gamma}$ move away from γ^{QR}; subsequently, comparable behavior to that in experiment 21 occurs.

In experiments 23 and 24, $\tilde{\gamma}^0$ is chosen to be a unit vector in the direction of minimum and maximum components of γ^{QR}, respectively, with a norm equal to that of γ^{ex}. In experiment 23, the inner products of $\tilde{\gamma}$ with γ^{ex} and γ^{QR} are both initially zero, and, similar to the case $\tilde{\gamma}^0 = 0$, the iterated solution moves toward γ^{ex}, with very little percentage change in $E_i{}^{QR}$ or increase in $|L_i{}^{QR}|$ from zero. In experiment 24, the behavior is essentially the same on all fronts as that in experiment 23, even though the unit vector $\tilde{\gamma}^0$ was chosen to maximize its orientation

towards γ^{QR}. The problem, of course, was that $||\tilde{\gamma}^0||$ was chosen far too small; it would be interesting to repeat these two experiments for $||\tilde{\gamma}^0|| = ||\gamma^{QR}||$.

To see the result of moving to the "other side" from zero of γ^{ex} for $\tilde{\gamma}^0$, $\tilde{\gamma}^0$ was chosen, in experiment 25, to be $2\gamma^{ex}$. The main difference from using $\tilde{\gamma}^0 = 0$ (experiment 7) is the sign reversal in the imaginary parts of L_i^{ex} and L_i^{QR}, indicating that indeed $\tilde{\gamma}$ is approaching γ^{ex} and γ^{QR} from the "other side." That is quite interesting, considering that for $\tilde{\gamma}^0 = 0$, $L_0^{ex} = 0$ and yet E_1^{ex} is close to E_1^{ex} for $\tilde{\gamma}^0 = 2\gamma^{ex}$, where $L_0^{ex} = 1.00$. This behavior is yet another indicator of the relative independence of the algorithm behavior and the inner products L_i.

One final experiment, number 26, used for $\tilde{\gamma}^0$ a uniform speed of sound throughout the object region, not equal to c_0 but to the average of c_0 and c_1. Also, the initial guess for the internal field is f^{inc}, in order to make this run more comparable to an actual experimental situation. (This is a low contrast run comparable to the higher contrast cases for the same type of experiment described in Section 5.7g.) The results after one iteration are nearly identical to the comparable experiment, experiment 9, where the initial speed of sound estimate for all pixels was c_0. By iteration 4, the results are even closer; an independence of reconstruction quality is demonstrated from the value of uniform speed of sound chosen for the object region initial guess, as noted earlier in Section 5.7g. This completes the discussion of the sequence of experiments, the results of which are displayed in Tables (5.14b.1a) through (5.14b.1d).

One last study was a variation in the degree of over/underdetermination quantified by the parameter Q_2, defined in Eq. (3.2g.2). A series of runs for an $n_{max} = 11$ grid reconstructed cylinders of size ka = 4.7 and speed of sound contrast $c_1 = 1.575$ mm/μsec for varying values of Q_2: 1.16, 0.98, 0.06, -0.47, and -0.74. These values of Q_2 correspond to 2.1, 2.0, 1.1, 0.5, and 0.26 times as many equations as unknowns. ($Q_2 < 0$ indicates underdetermination.) Note: In this study, only the number of transmitters was varied; the number of detectors was held constant. This was for two reasons: 1) the D_{mj} coefficients would have to be recomputed for each run, and mainly 2) the number of receivers would have to be varied in steps of eight, too coarse a gradation for this study. The squared errors in $\tilde{\gamma}$ are given in Table (5.14b.2) for 4 to 9 iterations of the sinc basis method. The initial guess for γ was zero, and the initial guess for the field was the incident field. More iterations were allowed for lower Q_2 to determine if convergence, though slower, actually would take place. A squared error of 180 is a rough threshold for acceptable reconstructions for this particular cylinder. Thus, after five iterations, $Q_2 = 0.06$ reconstructions are acceptable, but for $Q_2 < 0$, results are far from acceptable. Actually, as noted in Section 5.5, one could greatly reduce Q_2 if the receivers are clustered around the regions of largest scattered field energy, with comparable results in shorter execution times.

CHAPTER 6

CONCLUSIONS

To advance in high resolution ultrasonic imaging, two paths may be followed. One path, that of the perfection of B-scan systems, is actively being pursued by corporations currently involved in the production and marketing of such systems. Although quantitative distributions of physical ultrasonic parameters are not available from these echo images, high resolution, real time scanners with many features, are currently in use. The other path involves using the techniques of inverse scattering to ascertain quantitative estimations of the induced sources within the tissue of interest which result from the interaction of selected incident ultrasonic waves with the tissue inhomogeneities. As is evident from Chapter 2, the currently available first-order algorithms are not only often difficult to implement, but are severely restricted in the range of problems they can approximately solve. Consequently, although the quantitative images obtainable from these methods would be very useful in making the body of knowledge of tissue characterization more useful in medical diagnosis, there are no commercially available systems precisely because of these restrictions on range of applicability of the algorithms as well as their high computational complexity. One method with more promise than the first-order diffraction tomography algorithms discussed in Chapter 2 is the sinc basis moment method, discussed in the remaining chapters. Because of its iterative, coupled equations form, and its

fidelity to the exact integral equations (given the discretization/sinc basis expansion modifications), higher-order solutions are possible, and have been shown to be significant improvements over a typical first-order (Born approximation-based) inversion method reconstruction. However, the reconstructions obtained by the sinc basis moment method are again valid over only an unrealistically narrow range of problems. Although the order of complexity was reduced significantly (Chapter 4), it is still a very slow algorithm, and also has very large storage requirements. These limitations of storage and speed are, however, of secondary importance in the long run, with the advent of ever faster and larger computers. It is the physical limitation concerning phase shift addressed in Chapter 5 that needs immediate attention if the sinc basis method is ever to be seriously considered as a practical way of obtaining useful medical images. The basic area of signal processing of unwrapped phase signals (such as addition, with the constraint of commutativity) would be a good place to start. Because unwrapped phase signals are noise-sensitive and difficult to reliably obtain, even progress in the theoretical side is no guarantee of the possibility of developing a practical, working system. But even if a tomographic system using this method is not practical, perhaps other uses for this formulation may be helpful in areas of field computations such as in hyperthermia research, where field distributions coupled with the bioheat equation could conceivably result in predicted temperature distributions in a given insonified tissue.