

VARIATIONAL AND ACTIVE SURFACE
TECHNIQUES FOR ACOUSTIC AND
ELECTROMAGNETIC IMAGING

A Dissertation
Presented to
The Academic Faculty

By

Daniel A. Cook

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in
Electrical and Computer Engineering



School of Electrical and Computer Engineering
Georgia Institute of Technology
May 2015

Copyright © 2015 by Daniel A. Cook

VARIATIONAL AND ACTIVE SURFACE
TECHNIQUES FOR ACOUSTIC AND
ELECTROMAGNETIC IMAGING

Approved by:

Dr. Francesco Fedele, Advisor
*School of Civil and
Environmental Engineering
Georgia Institute of Technology*

Dr. Patricio Vela
*School of Electrical and
Computer Engineering
Georgia Institute of Technology*

Dr. Anthony Yezzi, Co-advisor
*School of Electrical and
Computer Engineering
Georgia Institute of Technology*

Dr. Karim Sabra
*School of Mechanical Engineering
Georgia Institute of Technology*

Dr. Jennifer Michaels
*School of Electrical and
Computer Engineering
Georgia Institute of Technology*

Dr. Alessio Medda
*Aerospace, Transportation, and
Advanced Systems Laboratory
Georgia Tech Research Institute*

Date Approved: March 23, 2015

*I have seen flowers come in stony places
And kind things done by men with ugly faces,
And the gold cup won by the worst horse at the races,
So I trust, too.*

–John Masefield, “An Epilogue”

To my parents and grandparents.

ACKNOWLEDGMENT

I gratefully acknowledge support provided by the Office of Naval Research Code 32 and by the Georgia Tech Research Institute. I am also appreciative of the longstanding collaboration with Roy Hansen and others at the Norwegian Defence Research Establishment, as well as my colleagues Tony Lyons and Dan Brown. In addition, the advice of Professors Jennifer and Thomas Michaels of the GT QUEST Laboratory was invaluable, as was the assistance of Alexander Dawson and Xin Chen.

I feel especially fortunate to have had Professors Francesco Fedele and Anthony Yezzi as co-advisors for this research. Their enthusiasm for their subjects is contagious, and their willingness and ability to communicate their knowledge to others is remarkable.

Lastly, I am indebted to my family, who supported this project with their enduring patience and understanding.

TABLE OF CONTENTS

ACKNOWLEDGMENT	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
SUMMARY	xi
CHAPTER 1 INTRODUCTION	1
1.1 Motivation	3
1.2 Variational Methods for Sonar and Radar Imaging	5
1.2.1 Radar and Sonar Imaging in the Presence of Multipath	6
1.2.2 Interferometric Phase Unwrapping	7
1.2.3 Synthetic Aperture Autofocus	10
1.2.4 Speckle Reduction	11
1.2.5 Anomalous Refraction in Synthetic Aperture Imagery	12
1.3 Nondestructive Testing and Evaluation	13
1.4 Adjoint Active Surfaces for Imaging	14
CHAPTER 2 CONNECTIONS AMONG SYNTHETIC APERTURE IMAGING TECHNIQUES	16
2.1 Survey of Imaging Modalities	17
2.1.1 Synthetic Aperture Radar	17
2.1.2 Synthetic Aperture Sonar	18
2.1.3 X-Ray Computed Tomography	19
2.1.4 Seismic Imaging	20
2.1.5 Radio Astronomy	21
2.2 Image Reconstruction	22
CHAPTER 3 ADJOINT ACTIVE SURFACES FOR LOCALIZA- TION AND IMAGING	25
3.1 The Adjoint Operator	26
3.2 The Green's Function	29
3.3 Adjoint Methods for Inverse Problems	30
3.4 Shape Gradient for Opaque Scatterers in an Unbounded Domain	33
3.5 Shape Gradient for Complex Scatterers in a Bounded Domain	37
CHAPTER 4 NUMERICAL EXPERIMENTS FOR ADJOINT AC- TIVE SURFACES	47
4.1 Closed-Form Solution for Forward and Adjoint Problems	47
4.2 Closed-Form Solution for Spherical Objects	52

CHAPTER 5	ADJOINT ACTIVE SURFACES FOR SCATTERING OF ULTRASONIC LAMB WAVES	57
5.1	Elastic Wave Propagation in Plates	58
5.2	Elastic Wave Scattering in Plates	61
5.3	Experimental Setup and Signal Processing	65
5.4	Model-Data Comparison	71
5.5	Shape Recovery via Active Surfaces	78
5.6	Shape Recovery via Nonlinear Optimization	85
5.6.1	Nonlinear Least Squares	85
5.6.2	Simplex Methods	88
CHAPTER 6	DETECTION OF REFRACTIVE ANOMALIES IN SAS IMAGERY USING OPTICAL FLOW	90
6.1	Breaking Solitary Waves	92
6.2	Optical Flow	94
6.3	Image Frame Generation	96
6.4	Experimental Approach	98
6.4.1	Weighting the Smoothness Constraint	98
6.4.2	Image Resolution	100
6.4.3	Spectral and Temporal Overlap	100
6.5	Experimental Results	101
6.5.1	Raytracing	102
6.5.2	Application to Sea Test Data	105
6.6	Discussion	107
CHAPTER 7	SUMMARY AND CONCLUSIONS	111
7.1	Open Questions and Suggested Future Work	113
REFERENCES		115

LIST OF TABLES

Table 3.1	Choices for boundary conditions on the interior surfaces Γ^j used to obtain the shape gradient.	37
Table 4.1	Spherical Hankel functions defined in terms of ordinary Bessel functions of the first and second kinds.	48
Table 5.1	Positions of the piezoelectric transducers bonded to the experimental plate.	66
Table 5.2	Phase speed and wavelength for the S_0 and A_0 modes at several transmitted frequencies.	69
Table 6.1	Parameters available for controlling optical flow.	98

LIST OF FIGURES

Figure 1.1 SAR intensity image and wrapped interferogram of Kilauea, collected by the C/X-Band Spaceborne Imaging Radar (SIR-C/X) SAR.	8
Figure 1.2 SAS images of the same scene collected on three different passes on the same day.	13
Figure 3.1 Source illuminating an inclusion embedded in a domain Ω^0 containing several measurement locations.	34
Figure 3.2 Source illuminating an inclusion embedded in a (possibly bounded) domain Ω^0 containing several measurement locations.	38
Figure 4.1 Spherical harmonics of degree n and order m	49
Figure 4.2 Coordinates used for the spherical harmonic expansion of u and v	50
Figure 4.3 Example gradient descent path for shape reconstruction.	52
Figure 4.4 Effect of noise on shape reconstruction.	55
Figure 4.5 Improving shape reconstruction by increasing measurement density.	55
Figure 5.1 Illustration and photo of components used in the ultrasonic plate experiment.	67
Figure 5.2 Phase and group velocities and the corresponding wavelengths for S_0 and A_0 Lamb wave modes.	68
Figure 5.3 Relative power spectral density for S_0 and A_0 modes.	70
Figure 5.4 Measured wavefields at six different frequencies within the transmitted bandwidth.	72
Figure 5.5 Slice through the measured wavefield, extracted along a radial line originating at Source 1 and passing through the center of the scan area.	73
Figure 5.6 Two-dimensional spectrum of the data extracted from the slice shown in Figure 5.5 filtered to extract only the desired mode of propagation.	74
Figure 5.7 S_0 and A_0 mode propagation along the radial line shown in Figure 5.5.	75
Figure 5.8 Comparison between the wavefield predicted using the Fourier expansion and the experimentally measured wavefield.	76

Figure 5.9 Comparison between the wavefield predicted using the Fourier expansion and the experimentally-measured wavefield.	77
Figure 5.10 Energy landscape for Lamb wave reconstruction at 10 kHz.	80
Figure 5.11 Energy landscape for Lamb wave reconstruction at 25 kHz.	81
Figure 5.12 Energy landscape for Lamb wave reconstruction at 50 kHz.	82
Figure 5.13 Energy landscape for Lamb wave reconstruction at 75 kHz.	83
Figure 5.14 Energy landscape for Lamb wave reconstruction at 100 kHz.	84
Figure 6.1 SAS images of the same scene collected on three different passes on the same day, demonstrating the presence of transient refractive effects.	91
Figure 6.2 Sand ripples observed in the Gulf of Mexico by the Office of Naval Research SAS12 synthetic aperture sonar.	91
Figure 6.3 Numerical simulation of the density field associated with a breaking soliton and the resulting bolus. (Figure from Venayagamoorthy and Fringer [107].)	93
Figure 6.4 Acoustic soundings were used to observe soliton breaking off of Ile-aux-Lièvres Island in the St. Lawrence Estuary in 2004. (Figure from Bourgault et al. [108].)	93
Figure 6.5 Rays traced through a hypothetical internal wave bolus using a geometry similar to the collection used in the at-sea testing.	102
Figure 6.6 Legend for HSV optical flow representation.	103
Figure 6.7 Optical flow for simulated SAS image.	104
Figure 6.8 Optical flow for isolated transient refractive effect.	108
Figure 6.9 Optical flow for complex refractive environment.	109

SUMMARY

This research seeks to expand the role of variational and adjoint processing methods into segments of the sonar, radar, and nondestructive testing communities where they have not yet been widely introduced. First, synthetic aperture reconstruction is expressed in terms of the adjoint operator. Many, if not all, practical imaging modalities can be traced back to this general result, as the adjoint is the foundation for backprojection-type algorithms.

Next, active surfaces are developed in the context of the Helmholtz equation for the cases of opaque scatterers (i.e., with no interior field) embedded in free space, and penetrable scatterers embedded in a volume which may be bounded. The latter are demonstrated numerically using closed-form solutions based on spherical harmonics. The former case was chosen as the basis for a laboratory experiment using Lamb waves in an aluminum plate. Lamb wave propagation in plates is accurately described by the Helmholtz equation, where the field quantity is the displacement potential. However, the boundary conditions associated with the displacement potential formulation of Lamb waves are incompatible with the shape gradient derived for the Helmholtz equation, except for very long or very short wavelengths.

Lastly, optical flow is used to solve a new and unique problem in the field of synthetic aperture sonar. Areas of acoustic focusing and dilution attributable to refraction can sometimes resemble the natural bathymetry of the ocean floor. The difference is often visually indistinguishable, so it is desirable to have a means of detecting these transient refractive effects without having to repeat the survey. Optical flow proved to be effective for this purpose, and it is shown that the parameters used to control the algorithm can be linked to known properties of the data collection and scattering physics.

CHAPTER 1

INTRODUCTION

Historically, the majority of signal and image processing methods are based on the notion of linearity and time invariance (LTI). These conditions open the door to powerful tools based on the Fourier transform such as the equivalence between convolution in the signal domain and multiplication in the frequency domain. When applications stray from being LTI, these techniques can be applied piecewise to subsets of the signal for which LTI is assumed to hold. Meanwhile, recent years have witnessed the maturity of a different class of processing techniques based on partial differential equations (PDEs). Its strength is found in the inherent ability of PDEs to respond to very localized signal conditions. The notion of short-time, or windowed, LTI processing is unnecessary.

PDE-based processing schemes are often derived using the calculus of variations which is aimed at finding an extremal function that satisfies a given optimality condition. The optimality condition is stated in integral form and can incorporate a rich set of constraints. Two common types are (1) smoothness constraints, and (2) fidelity constraints which impose a certain governing model on the underlying signal. The former are akin to least-squares in that their purpose is to discourage unrealistic departures from the expected model and/or to mitigate the effect of measurement errors. The latter type of constraint is extremely powerful, as it allows the user to introduce the actual physical model governing the measured signal or to introduce artificial models whose purpose is to elicit the desired response from the data.

The resulting PDEs do not typically permit closed-form solutions, and signal processing using PDE techniques inevitably requires some level of familiarity with numerical solution schemes such as finite differences, finite elements, or boundary elements. The process of finding the extremal function sometimes exposes an operator

known as the adjoint which, while purely a mathematical construct, is often associated with projecting measurements backward through the system that generated them from the source. This notion of backprojection, or backpropagation, is extremely powerful and appears throughout the inverse problem literature as well as other fields of study.

The research contained in this dissertation is motivated by a desire to expand the role of variational and adjoint processing methods into segments of the sensing and imaging communities where they have not yet been widely introduced. This is accomplished through the following contributions:

1. The connections among several imaging modes are explored. Discussing disparate applications in a common framework helps to lower the barriers to exchanging ideas across disciplines.
2. Adjoint active surfaces are developed for physical inverse problems governed by the Helmholtz equation. This approach is then applied to experimental data using ultrasonic elastic waves to locate damage in a thin plate.
3. Optical flow techniques are applied to synthetic aperture sonar imagery to detect the presence of internal wave-like boluses of dense water moving over the sea floor. To date, this is the most successful method for detecting these refractive anomalies without resorting to repeat-pass imaging.

Several publications have been written in the pursuit of these objectives:

- T. M. Benson, D. P. Campbell, D. A. Cook, “Gigapixel spotlight synthetic aperture radar backprojection using clusters of GPUs and CUDA,” in Proceedings of the 2012 IEEE Radar Conference, pp. 853-858.
- D. A. Cook, “Chapter 7: Spotlight Mode SAR,” in W. M. Melvin and J. A. Scheer, eds., *Principles of Modern Radar: Vol. II Advanced Techniques*. SciTech Publishing, Inc., 2013.

- D. Cook, D. Brown, and Z. Lowe, “Synthetic aperture sonar contrast (Invited Paper),” in *1st International Conference and Exhibition on Underwater Acoustics*, June, 2013, pp. 143–150.
- R. E. Hansen, A. P. Lyons, T. Sæbø, H. J. Callow, and D. A. Cook, “The effect of internal wave-related features on synthetic aperture sonar,” *IEEE Journal of Oceanic Engineering* (Accepted for publication and in press), 2014.
- D. A. Cook, “Connections among synthetic aperture imaging techniques,” Invited Paper, *10th European Conference on Synthetic Aperture Radar (EUSAR 2014)*.
- D. C. Brown, D. A. Cook, and A. P. Lyons, “Spatial coherence theory and its application to synthetic aperture systems,” in *Proceedings of the International Conference on Synthetic Aperture Sonar and Synthetic Aperture Radar (ICSASSAR)*, Institute of Acoustics, 2014.
- D. A. Cook, R. E. Hansen, A. P. Lyons, and A. J. Yezzi, “Motion tracking of transient refractive effects in SAS imagery using optical flow,” in *Proceedings of the International Conference on Synthetic Aperture Sonar and Synthetic Aperture Radar (ICSASSAR)*, Institute of Acoustics, 2014.
- D. A. Cook, M. Mueller, F. Fedele, and A. J. Yezzi, “Adjoint active surfaces for localization and imaging,” *IEEE Transactions on Image Processing*, pp. 316–331, vol. 24, no. 1, 2015.

1.1 Motivation

Traditional imaging techniques use a set of measurements, generally of scattered or absorbed radiation, to estimate a desired property within a volume. The property

being mapped (density, reflectivity, absorption, etc.) conveys, either directly or indirectly, useful information about the volume being interrogated. For example, synthetic aperture radar [1, 2, 3] and sonar [4, 5, 6] are used to create reflectivity maps of the terrain illuminated by the sensor. Seismic oceanography uses low frequency (1–200 Hz) sound that is emitted from air guns, reflected from oceanic structures, and received on a towed array of hydrophones to image the water column. Seismic imaging is also used for subsoil mechanical characterization, an important issue in many fields of both Earth sciences and geotechnical engineering [7, 8, 9]. Surface wave methods utilize the dispersive nature of seismic waves in a heterogeneous medium to obtain shear velocities and infer the lateral variations of subsurface layers [10, 11]. In the field of medical imaging, X-ray tomography maps the absorption characteristics of tissue within the body [12, 13, 14]. Physicians are able to diagnose conditions accompanied by changes in absorption relative to what would be observed under normal healthy circumstances. Similar problems are encountered in nondestructive testing and evaluation and structural health monitoring, where the goal is often to assess the integrity of a structure using few measurements that may be located far away from the region of interest.

Variational methods for image processing and computer vision have provided successful advancements for a number of applications. Interestingly, these advances have not diffused uniformly across disciplines. For example, variational methods were embraced early on by the medical imaging community. The precise segmentations offered by level sets and active contours are useful for obtaining the quantitative measurements which are vital to effective diagnosis and treatment. Meanwhile, other disciplines have had scant contact with variational methods. Synthetic aperture radar and sonar are good examples of this. On one hand, there is little need for significant changes to current image reconstruction techniques and processing architectures because they work well for a large fraction of the situations that are of practical interest.

On the other, the remaining cases tend to be very challenging and are in need of real innovation. Several applications illustrate this point, and a subset of these is the focus of this research.

1.2 Variational Methods for Sonar and Radar Imaging

Traditional SAR and SAS image reconstruction is based on the following set of assumptions [15]:

1. The scene can be modeled as a continuum of infinitesimally small point scatterers whose reflections do not affect one another. This is known in mathematical physics as the Born approximation [16]. In other words, it is assumed that there is no multiple scattering and that the principle of superposition applies to the reflected signals. This linearity of the signals is one reason it is possible to reconstruct accurate imagery without prior knowledge of the scene content.
2. The incident radiation does not penetrate the reflecting surface. Consequently, the scene can be modeled as a sheet which is not necessarily flat, but is infinitesimally thin. It is this model that gives meaning to the interferometric applications of SAR (SAS) used for deriving topographic (bathymetric) information from image pairs. When the scene deviates vertically from a plane, the resulting image has a phase component which can be used to infer the terrain height [17].
3. The medium is assumed to be infinite in extent, so there are no boundaries to consider. This permits the use of the free-space Green's function for the reconstruction. The presence of boundaries can give rise to multipath effects resulting in multiple reflections, or ghosts, in imagery. This effect is particularly evident in SAR imagery of urban areas and SAS imagery collected in shallow water.

4. The scene content does not move during the data collection interval. The relative motion between sensor and scene is the phenomenon used to resolve and place scene content in the cross-range dimension. An object moving against the static background may not appear in the imagery at the correct location. A classic example often cited is the phenomenon of imaging a moving train. If the train's velocity vector has any component along the radar's line of sight, the cars will appear displaced in cross-range. It is possible to get rather disconcerting SAR imagery showing a well-focused train that appears to be riding off its track.

These assumptions are somewhat predictable, representing a commonsense approach to modeling a simplified environment. They also have deeper mathematical significance, enabling the development of simple solutions to the wave equation governing the propagation of electromagnetic and acoustic waves. In particular, restrictions (1) and (3) give rise to the term $e^{-ikR}/4\pi R$, which is ubiquitous in the radar and sonar disciplines. This is the Green's function representing the solution to the free space wave equation [18].

1.2.1 Radar and Sonar Imaging in the Presence of Multipath

The first problem illustrating the need for advancement beyond the standard approaches is that of visualizing the interior structure of buildings using radio waves. This is known in the literature as through-the-wall-imaging (TWI). The task is made difficult by the fact that the electromagnetic properties of the walls are unknown and multiple reflections within the structure tend to corrupt the imagery. Broadly speaking, the TWI literature falls into two categories. The first approaches the problem by expanding, or amending, classical beamforming and reconstruction algorithms in an attempt to produce an image product with fewer artifacts [19, 20]. For example, [20]

proposes the use of Snell’s law to compensate for the RF propagation through a single homogeneous wall. The wall imparts an additional delay relative to that required for free-space beamforming that tends to defocus the imagery unless it is accounted for. While improving focus, this technique does not address the problem of multipath returns that would be present in many structures [21]. Another common approach to TWI is to remove image artifacts by using autofocus, deconvolution, and postprocessing techniques inspired by the SAR, geophysics, and radio astronomy communities. In particular, [22] and [23] make use of the CLEAN algorithm [24] which has enjoyed success in many applications where there is the need to recover an impulse response that has been corrupted by nonideal behavior in the acquisition system.

Another application in which traditional processing methods fail is shallow-water acoustic imaging using synthetic aperture sonar. In deep water, the sonar transmits acoustic signals (or pings) and measures the reflection from the sea floor alone. In shallow water, energy that radiates upward can reflect directly from the sea surface, or it can bounce between the bottom and surface before being received. The result is a corrupted image of the sea floor. Imaging in a shallow water multipath environment is arguably more difficult than the TWI problem because the sea surface moves during the data acquisition. The traditional method of mitigating this problem is to control the transmit and receive beam patterns of the sensor so as to place as little gain as possible in the direction of the surface. Adaptive beamforming methods have also been considered [25]. It seems possible that variational methods could be applied under certain conditions. This begs the broader question of the nature of the connections between traditional adaptive beamforming methods [26] and possible variational approaches.

1.2.2 Interferometric Phase Unwrapping

When synthetic aperture images are collected along parallel paths, with a slight difference in grazing angle, the resulting image pair can be used to estimate the terrain

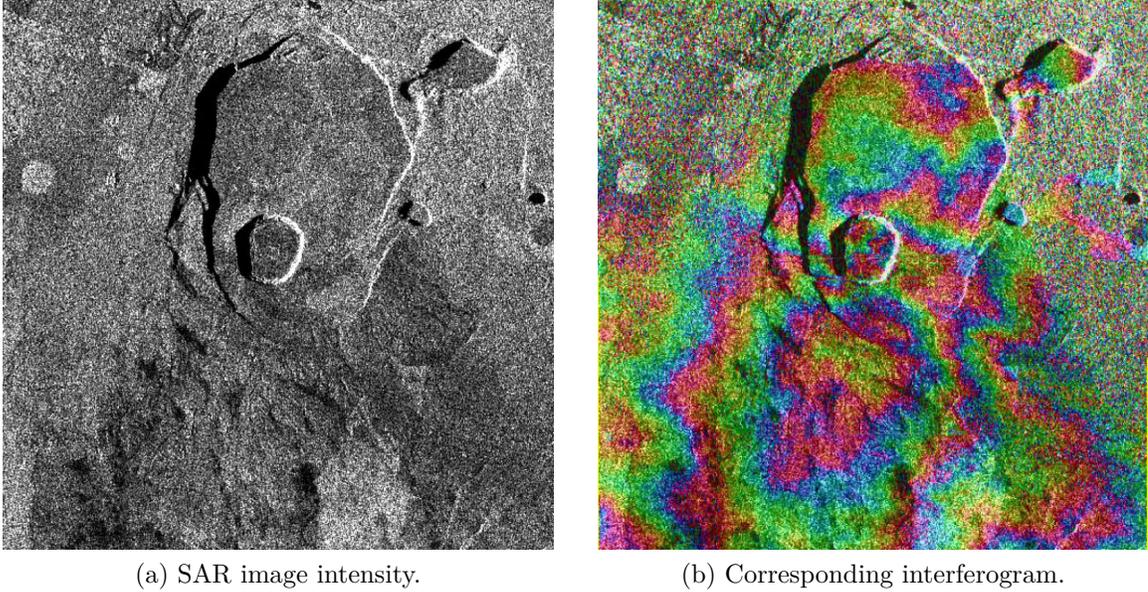


Figure 1.1: SAR amplitude image and wrapped interferogram of Kilauea, collected by the C/X-Band Spaceborne Imaging Radar (SIR-C/X) SAR. Images courtesy NASA/JPL-Caltech.

height [3, 27, 17]. The complex-valued images are first co-registered and the phases are then compared on a pointwise basis. The phase difference is related to the terrain height. Since typical imaging radar and sonar wavelengths range between approximately 1–30 cm, it is nearly always the case that the terrain height variations within the scene span multiple wavelengths. This situation creates ambiguities that must be resolved through a suitable two-dimensional phase unwrapping procedure. An example of a wrapped SAR interferogram is shown in Figure 1.1.

Phase unwrapping is one of two problems in the fields of SAR and SAS that have experienced significant contribution from variational processing techniques. This work was described in a series of papers by Ghiglia and culminated in a chapter in a 1998 book devoted to two-dimensional phase unwrapping [28]. Unfortunately, there does not appear to have been much research in this area since. The variational approach to phase unwrapping is outlined below.

The object is to find the 2D phase function $\phi(x, y)$ that minimizes the cost functional

$$J = \int \int f(\phi_x, \phi_y, x, y) dx dy, \quad (1.1)$$

where $f = |\phi_x - \psi_x|^p + |\phi_y - \psi_y|^p$, and $\psi(x, y)$ is the wrapped phase function derived from the co-registered interferometric image pair. The resulting Euler-Lagrange equation is

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial \phi_x} \right) + \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial \phi_y} \right) = 0. \quad (1.2)$$

Substituting the partial derivatives of f with respect to ϕ_x and ϕ_y gives the following differential equation for ϕ :

$$\frac{\partial}{\partial x} [U(x, y)(\phi_x - \psi_x)] + \frac{\partial}{\partial y} [V(x, y)(\phi_y - \psi_y)] = 0, \quad (1.3)$$

where

$$\begin{aligned} U(x, y) &= |\phi_x - \psi_x|^{p-2} \\ V(x, y) &= |\phi_y - \psi_y|^{p-2}. \end{aligned} \quad (1.4)$$

The functions U and V can be thought of as generalized data-dependent weighting functions on the phase derivatives. These should not be confused with optional external weights which might be supplied to incorporate prior knowledge of the spatially-varying data quality. When $p = 2$, U and V are both equal to one, giving Poisson's equation

$$\phi_{xx} + \phi_{yy} = \psi_{xx} + \psi_{yy}, \quad (1.5)$$

or simply $\nabla^2 \phi = \rho$, where $\rho = \psi_{xx} + \psi_{yy}$ captures the measured data. The unweighted least squares 2D phase unwrapping problem is thus reduced to Poisson's equation, and it may be solved using any of a number of efficient numerical schemes. If the weights U and V are independent of ϕ , the result is a weighted least squares problem that can be solved using a different set of numerical approaches. This latter case is the most typical, as it is common in practice to assign weights of zero to portions of

the data known to be of poor quality. Such would be the case in regions of shadow, where the SNR is very low.

1.2.3 Synthetic Aperture Autofocus

The creation of synthetic aperture imagery is highly dependent on having a well-sampled spatial aperture. For radio astronomy the difficulty is collecting a data set with adequate support in the Fourier domain. When the support is irregular or incomplete, the resulting imagery is likely to contain strong artifacts [29]. The standard method for removing these artifacts is based on the CLEAN algorithm [24]. For SAR and SAS imaging, the problem is not that of achieving complete support in the Fourier domain. Rather, the challenge is collecting data along a known path. While radio telescopes use the Earth’s steady rotation for aperture synthesis, SAR and SAS sensors are carried by aircraft and submersibles that move through fluid media. It is difficult to travel along an ideal straight-line trajectory, yet deviations larger than $\lambda/8$ can cause unwanted image artifacts. Effective motion measurement and compensation, rather than corrective postprocessing, is generally the preferred solution [5].

Autofocus methods were developed to remove residual motion and hardware-induced errors from the imagery without relying on other sensor measurements. There are a number of successful techniques that have become standard for spotlight-mode SAR [2, 3]. For various reasons, SAS autofocus is a more difficult problem, and the state of the art remains somewhat in flux.

There is a family of SAR and SAS autofocus techniques that resembles variational methods [30, 31, 32], although the term ‘variational’ does not seem to have been used in the literature. These begin by choosing a scalar-valued metric, such as contrast or sharpness [33], to represent the quality of the entire image. The gradient of this metric is then taken with respect to an unknown one-dimensional error function $\phi(x)$, and a gradient descent optimization is performed to find the particular $\phi(x)$ that maximizes

the image quality according to the chosen metric. It is important to point out that there does not exist a universally-good image quality metric, as image quality is inherently subjective and dependent on the scene content.

These variational autofocus techniques must operate on a form of the raw sensor data, while the objective metric is computed from the resulting image. This means that some level of processing is required at each iteration of the algorithm to transform to the image domain. Unfortunately, synthetic aperture image reconstruction can be quite expensive. This problem is bypassed and computational efficiency achieved by assuming that the magnitude of the function $\phi(x)$ is bounded by the range resolution of the sensor. This restriction permits the user to operate on a data set that is closer to the state of the final image, thus reducing the computational cost of each iteration.

While autofocus is not part of this research, the topic appears to be ripe for a methodical treatment in the literature using a variational framework. Additionally, given the computing power that is commonly available today, the restrictions placed on $\phi(x)$ should be revisited. For example, the prevalence of extremely powerful commodity GPUs is enabling the use of simple, robust, brute force reconstruction techniques, obviating the need for many of the efficient, but approximate, methods developed by the SAR and SAS communities through the beginning of the 21st century [34].

1.2.4 Speckle Reduction

Speckle is a characteristic phenomenon of coherent imaging systems, be they optical, microwave, or acoustic. Although often referred to colloquially as ‘speckle noise,’ it is not a form of noise in the true sense of the word. Speckle arises when the surface being imaged is rough compared to the wavelength of the incident energy. Under these conditions, the signal observed within any given resolution cell is the phasor sum of many randomly distributed scattering centers. It has been shown that the probability density function of speckle intensity, I , is described by the negative exponential distribution [35]. An important property of this distribution is that the

standard deviation is equal to the mean. Furthermore, that the intensity fluctuation is comparable to its mean value is evidence of the fact that speckle can be visually distracting and can mask details within the imagery. The average image contrast K in speckle-dominated regions is equal to unity when it is defined as $K = \sigma_I/\mu_I$.

Within the fields of synthetic aperture sonar and radar, speckle reduction has probably received the most attention from researchers wishing to apply variational and PDE-based image processing techniques. The motivation is strong because speckle is ubiquitous and can affect image interpretation. Furthermore, the traditional method of speckle reduction, known as multilooking, mitigates speckle at the cost of image resolution, which is a precious commodity indeed. PDE-based methods are attractive because they possess localized behavior driven by the scene content [36, 37]. This property results in an ability to preserve edges. Consequently, anisotropic diffusion techniques have been applied to the problem of locating linear features, namely pipelines and undersea cables, in SAS imagery [38, 39].

1.2.5 Anomalous Refraction in Synthetic Aperture Imagery

Traveling boluses induced by breaking internal waves have recently been observed in synthetic aperture sonar imagery, where their appearance resembles natural seabed ripple patterns [40]. These boluses cause a lensing effect that concentrates the acoustic energy in a manner resulting in regions of increased intensity followed downrange by reduced intensity. These were discovered because the same area was surveyed twice, and the apparent sand ripples shifted in a manner that was inconsistent with known sediment transport behavior in the area. A similar effect is well-known in the SAR community, where certain atmospheric conditions can produce ‘moving sand dunes’ in the imagery [41].

Example SAS imagery showing the refractive effect of these boluses is given in Figure 1.2. Detection is a difficult prospect and is of limited utility if repeat surveys are required. This provides motivation for finding detection methods that can be used

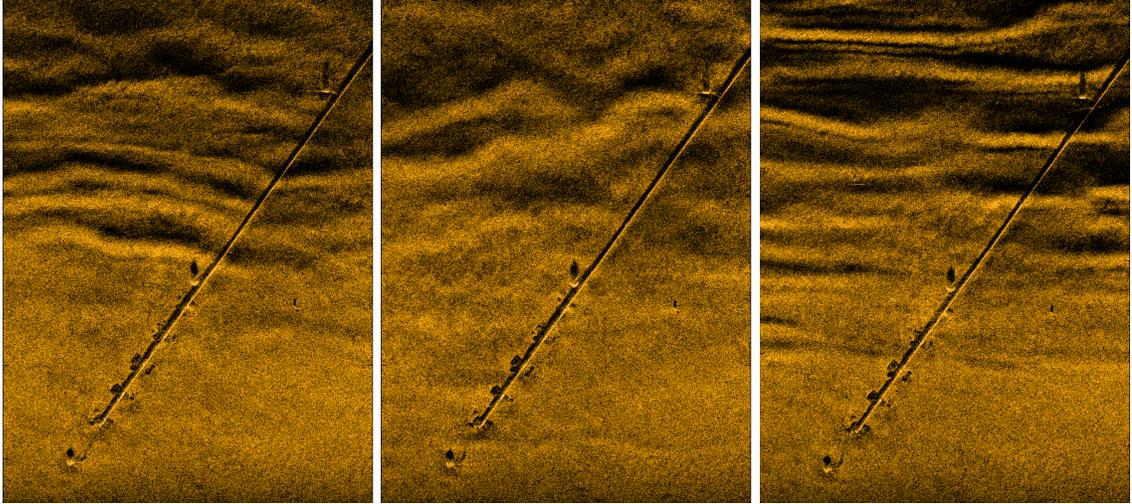


Figure 1.2: SAS images of the same scene were collected on three different passes on the same day, at times 07:19 (left), 08:18 (center), and 12:05 (right). The scene size is 80 m x 100 m and shows a pipeline with anchors on the seabed. The data were collected by FFI’s HUGIN AUV carrying an interferometric SAS during the ARISE12 trials onboard the NATO research vessel *Alliance*, outside Elba island, Italy.

on single-pass imagery. Some approaches are suggested in [42], and one of them is pursued in this research. Specifically, the use of optical flow for detecting the presence of these effects is suggested. Conventional speckle-tracking techniques (such as are used for interferometric image registration) are not suited to this problem because the individual image frames will have different speckle realizations.

1.3 Nondestructive Testing and Evaluation

Certain types of nondestructive testing bear a strong resemblance to sonar and radar array sensing applications, including synthetic aperture. This occurs in situations where the measurement locations are constrained to be relatively far removed from the area being interrogated and where fine resolution is desired. This is most obviously true for SAFT (synthetic aperture focusing technique) [43, 44]. A good deal of success has been realized by applying radar and sonar array processing techniques to NDE problems. In particular, the use of linearly-swept frequency modulated (LFM,

or chirp) waveforms has enabled the rapid acquisition of broadband signals [45]. Furthermore, space-time adaptive processing (STAP) techniques have recently shown their potential utility [46].

In spite of these similarities and examples of successful cross-discipline technology transfer, the typical NDE problem is, generally speaking, fundamentally more complicated than typical radar or sonar applications. The reason for this is simple: As previously discussed, radar and sonar usually assume free-space wave propagation, very nearly allowing the user to ‘forget about the physics’ in a great number of important applications. Meanwhile, ultrasonic NDE applications often involve one or more types of elastic wave propagation which are usually dispersive in nature [47]. Furthermore, these modes of propagation can couple into one another when waves interact with damage or the external boundaries of the solid [48].

Variational methods have been applied to NDE problems [49], but the applications are primarily for electromagnetic sensing modes, not ultrasonic. While the literature contains a number of theoretical treatments, there is a relative dearth of experimental results. One goal of the present research is to remedy this situation by applying adjoint image reconstruction techniques to a set of ultrasonic wave measurements collected in a laboratory.

1.4 Adjoint Active Surfaces for Imaging

In typical image processing applications active contours (surfaces) evolve in response to image pixel (voxel) intensity to find the boundary of a region of interest. A distinct advantage may be realized when the underlying measurements, or raw data, are available. The physical laws governing the wave propagation and the observed measurements are incorporated into the image reconstruction using the adjoint of the forward model. It is thus applicable whenever a suitable physical model can be identified and scattered or absorbed radiation can be observed. This differs from the

general trends in the imaging literature in at least two key respects. First, it is aimed at directly representing the geometry of the objects and boundaries of interest, as opposed to making a 2D or 3D map of reflectivity or absorption. A similar approach has been applied in the areas of medical optical tomography and RF underground imaging [50, 51, 52] with promising results. Secondly, the boundaries are included in the differential equations that govern the wave propagation. The technique should therefore be capable of good performance even in multipath environments.

CHAPTER 2

CONNECTIONS AMONG SYNTHETIC APERTURE IMAGING TECHNIQUES

It is common knowledge that the disciplines of synthetic aperture radar and sonar have exchanged technological advances for decades. Furthermore, both have benefited from developments in the seismic imaging community. Such advances have occurred only sporadically as researchers have become aware of cross-discipline similarities that are not always obvious. This chapter surveys several key synthetic aperture imaging fields and discusses fundamental commonalities among them. In particular, synthetic aperture radar and sonar, X-ray computed tomography, seismic imaging, and radio astronomy are considered.

Interestingly, synthetic aperture radar and sonar belong to a rather distinguished family of Fourier imaging techniques that has garnered several Nobel Prizes:

- X-ray crystallography (Max von Laue, 1914, Physics)
- Holography (Dennis Gabor, 1971, Physics)
- Radio astronomy (Sir Martin Ryle and Antony Hewish, 1974, Physics)
- X-ray tomography (Allan Cormack and Godfrey Hounsfield, 1979, Physiology or Medicine)
- Magnetic resonance imaging (Paul Lauterbur and Peter Mansfeld, 2003, Physiology or Medicine)

Synthetic aperture imaging practitioners have long recognized the existence of similarities between these disciplines and their own, and with the abundance of researchers, literature, and commercial products, it would seem as though ideas could flow freely

among related technological specialties. The barrier to information exchange is, however, quite high in practice for three reasons. First, researchers tend to spend their entire career specializing in a single field. Second, the scientific literature is permeated by domain-specific jargon that is difficult for the uninitiated to negotiate. Thirdly, the governing physics differ for each application. What might be a fundamental limitation for one imaging mode could be a trivial concern for another. This chapter begins to address the second and last of these concerns in hopes of creating an interest in deliberately working across disciplines to advance the collective state of the art.

2.1 Survey of Imaging Modalities

The previously-stated goal is approached by summarizing the key technical challenges associated with several imaging techniques.

2.1.1 Synthetic Aperture Radar

It is difficult to succinctly characterize the variety of SAR instruments and applications in existence. Broadly speaking, SAR operates at ranges on the order of 1–1000 km, and typical frequencies used for imaging range between 4 and 35 GHz (C band to Ka band). Image sizes are on the order of 1–10 km, and fine resolution is generally understood to be around 0.3 m or better.

SAR collections fall into two categories, stripmap and spotlight [1, 2, 3]. Stripmap mode provides high area coverage rates, but it exhibits relatively coarse cross-range resolution because the integration angle is limited to the antenna beamwidth. This problem is overcome by spotlight mode imaging, in which the beam is continually slewed to point toward a fixed location on the ground during the collection. The integration angle may then be much larger than the beamwidth. The beamwidth therefore limits the scene size of a spotlight image, as opposed to its resolution.

Since its inception in the 1950s by Carl Wiley, SAR technology has had to overcome a number of technological obstacles. Notable examples are the transition from

film recording and optical reconstruction to digital systems and the use of stretch receivers to overcome sampling rate limitations when employing large RF bandwidths. Today, there are no significant challenges to achieving the aforementioned fine resolution criteria of 0.3 m. Substantially improving upon this is a difficult task, however. For a given fractional bandwidth, increasing the transmitted RF bandwidth can be done by employing a higher center frequency. However, at high frequencies it becomes difficult to manufacture antennas and RF components capable of withstanding the power levels necessary for achieving good SNR.

2.1.2 Synthetic Aperture Sonar

Imaging underwater over any appreciable distance can only be achieved using sound waves, for light and radio waves attenuate far too quickly. Even high-frequency sound waves do not propagate very far. Fine resolution real-beam imaging sonars have limited ranges for this reason, and synthetic aperture sonar is the preferred choice for creating very detailed images of the sea floor. SAS typically operates somewhere in the band of 20–300 kHz. Typical range resolution is on the order of 3 cm, which is easy to achieve given current transducer technology and the fact that this resolution translates into only 25 kHz of transmitted bandwidth.

Two key innovations have made SAS a viable technology. The first is the use of the Vernier array, in which multiple receivers are used to increase the area coverage rate by allowing the sonar to extend its reception time [53, 54]. The increase in coverage rate is proportional to the number of receivers used. The second enabling technology is the use of redundant phase centers, in which a subset of the Vernier array is overlapped between successive pings [55]. The signals from the overlapping channels are cross-correlated, and the measured delay is used to infer the platform trajectory.

The most challenging problems for SAS come from operating in the extremes of ocean depth. Perhaps surprisingly, imaging in shallow water is very difficult. The

environment is not free space, as the sea surface acts as a constantly-moving mirror to produce multipath reflections that corrupt the data. Furthermore, shallow water may exhibit dramatic salinity and temperature gradients that cause refraction of the sound, whereas most reconstruction algorithms assume straight-line propagation. At the opposite end of the scale, it is difficult to manufacture vehicles and sensors that can travel to the deepest parts of the ocean. In addition to withstanding the extreme pressure, the unmanned vehicles that carry SAS must have enough on-board power to make the trip to and from the ocean floor while also having enough survey time to make the journey worthwhile. Acoustically speaking, however, operating in the deep ocean is relatively easy. The environment is free of surface reflections and the water is generally homogeneous, meaning that the actual propagation very closely matches the simple theoretical model used to derive reconstruction algorithms.

2.1.3 X-Ray Computed Tomography

X-ray CT is widely used for imaging the human body. It differs from SAR and SAS in that it measures transmitted, rather than reflected, energy. Regardless, the underlying mathematics of image reconstruction are remarkably similar to spotlight mode SAR [56, 57]. Many of the challenges of creating a good CT image involve trade-offs associated with improving the data quality at the expense of exposing the patient to an increased dosage of radiation [58]. Current medical CT scanners offer a resolution limit on the order of less than 1 mm. Manufacturing smaller detectors would improve image resolution, but would also be more susceptible to noise, necessitating a higher source intensity and increased risk to the patient. Facets of this balance between source flux and image quality shape many areas of research and vendor intellectual property.

Apart from the hardware itself, another key component of successful CT imaging is the ability to calibrate the data. Within the industry, this process is known as physical correction and is quite complicated because of the sophistication and sensitivity of

the equipment. In fact, the physical correction constitutes a significant fraction of the intellectual property associated with a commercial CT scanner.

Unwanted motion is another difficulty for CT reconstruction. Unlike SAR and SAS, the motion of the CT scanner is controlled and well-understood as it rotates around the patient. However, the motion of the body through breathing and heartbeat can easily corrupt the image. Tracking and correcting for this motion is an active area of research in the medical imaging community. Another serious problem for CT imaging is the presence of metal in the body, such as dental fillings and titanium replacement joints. These objects absorb the incident X-rays, effectively creating holes in the data that result in visual artifacts. Current research efforts are investigating methods of mitigating these artifacts and interpolating through the regions of missing data.

2.1.4 Seismic Imaging

Of all disciplines using synthetic aperture techniques, seismic imaging is probably the most mathematically sophisticated as well as the most difficult for outsiders to understand. The sophistication comes from the fact that the Earth's interior is a complex environment and is difficult to model, much less reconstruct from measurements. The steep learning curve results from the somewhat unusual circumstance that seismic imaging straddles pure science as well as heavy industry. The literature therefore combines deep concepts from physics and mathematics with a century's worth of petroleum field engineering jargon [59].

While other fields emphasize increasingly higher image fidelity with respect to the true scene properties (for example reflectivity or attenuation coefficient), the geophysical community is primarily interested in identifying the structure of reflecting layers within the earth. It is less concerned with solving for the exact physical properties that cause the reflections. Furthermore, certain environments may favor certain reconstruction techniques. In contrast, other disciplines have a clearer path toward a

single approach that will render the best possible image.

Some of the reconstruction methods used by the geophysics community overlap with other fields. In particular, what the SAR and SAS communities refer to as time-domain reconstruction, or backprojection, is known as Kirchhoff migration to geophysicists. Perhaps more significant is the ω - k algorithm developed by Stolt [60, 61]. It was a revolutionary step in that it achieves a mathematically exact reconstruction while employing the efficiency of the fast Fourier transform (FFT). It assumes a constant wave propagation speed, which is usually not a good fit to the subterranean environment. It is, however, an excellent model in the majority of radar and sonar applications. The ω - k algorithm is therefore a significant gift from the seismic community.

2.1.5 Radio Astronomy

Aperture synthesis for radio astronomy is an impressive technology, if for no other reasons than the sheer scales involved and the sensitivity of the equipment used to receive signals from deep space. While the aforementioned collection methods involve forming a synthetic aperture using a moving aircraft or a rotating X-ray detector, radio astronomy forms apertures using the Earth's rotation [29]. Of the techniques discussed here, radio astronomy is the only one that is passive. Images are formed on the principle that signals from pairs of antennas, called baselines, can be correlated and the degree of coherence measured as the Earth rotates. The rotation of the baseline separation vectors determines the loci of the sampled data, resulting in a two-dimensional coherence, or visibility, map.

The Van Cittert–Zernike theorem from statistical optics states that the visibility is related to the image intensity through the Fourier transform. Radio astronomy image reconstruction therefore strongly resembles that of spotlight SAR via the polar formatting algorithm. In both cases, the measured data can be treated as samples within the frequency domain representation of the image sought. The sampling is

usually nonlinear, so reconstruction involves interpolating the available data onto a rectangular grid of samples that is amenable to inversion via the FFT.

Radio astronomy spectra are also usually poorly sampled in some respect. This situation resulted in algorithms, such as CLEAN [24], that are used to improve the impulse response of the image. Some of these have migrated to the SAR and SAS communities. Geophysicists also use similar techniques for mitigating the effect of multiple reflections inside the earth.

2.2 Image Reconstruction

This section illustrates how reconstruction algorithms for the technologies listed above can all be traced back to a core set of mathematical principles, namely the adjoint and the pseudoinverse. Borrowing the notion of ‘exploding reflectors’ from the seismic literature, the scene is represented as a continuum of sources all radiating simultaneously, combining the reflectivity and transmitted signal into a single function $f(x, t)$. Reflection data can then be treated as though the signals originate at the reflectors, but travel with half of the actual propagation speed. This general model allows one to ignore the reflection/transmission nature of the specific problem. A linear operator T is next introduced that maps the virtual source field $f(x, t)$ into measurements $s(x_m, t)$ made at the set of observation points $\{x_m\}$. The mapping T is usually taken to be the convolution of $f(x, t)$ with the free space Green’s function. The frequency domain version (usually associated with the Helmholtz equation) will be preferred for the derivations to follow:

$$s(x_m, \omega) = T f(x', \omega) = \int_X \frac{e^{-i2kR}}{4\pi R} f(x', \omega) dx', \quad (2.1)$$

where $R = |x' - x_m|$ and $k = \omega/c$.

The adjoint operator T^* associated with T is defined by the relationship $\langle T^* s, f \rangle = \langle s, T f \rangle$, where the angle brackets denote the inner product (also see section 3.1). The adjoint can be found by substituting the forward mapping into this definition and

rearranging the result:

$$\begin{aligned}
\langle s, Tf \rangle &= \int_{\Omega} s^*(x_m, \omega) \int_X f(x', \omega) \frac{e^{-i2kR}}{4\pi R} dx' d\omega \\
&= \int_X f(x', \omega) \int_{\Omega} s^*(x_m, \omega) \frac{e^{-i2kR}}{4\pi R} d\omega dx' \\
&= \int_X f(x', \omega) \int_{\Omega} \left\{ s(x_m, \omega) \frac{e^{i2kR}}{4\pi R} \right\}^* d\omega dx' \\
&= \langle T^* s, f \rangle,
\end{aligned} \tag{2.2}$$

where the adjoint operator, applied to $s(x_m, \omega)$, is seen to be

$$T^* s(x_m, \omega) = \int_{\Omega} \frac{e^{i2kR}}{4\pi R} s(x_m, \omega) d\omega. \tag{2.3}$$

The operator T^* is identified as the backprojection operator since it maps points s in the measurement space back into the space of the scene. The corresponding reconstruction algorithm pseudocode is as follows:

```

Initialize the output image
for Each point  $x$  in the reconstructed image do
  for Each measured signal in the set  $\{s(x_m, \omega)\}$  do
    · Backproject all frequencies  $\omega$  from  $x_m$  to  $x$  using the adjoint operator  $T^*s(x_m, \omega)$ ;
    · Accumulate the resulting scalar value into the output image at location  $x$ ;
  end for
end for

```

While the adjoint operator (3.43) backprojects the measured data as desired, it does not compensate for effects such as spherical spreading. In fact, it introduces another spreading term. A filtered backprojection solution based on the pseudoinverse performs the desired compensation and also serves as the least squares solution when the measurements are subject to noise. The least squares solution is given by $f = (T^*T)^+ T^*s$, where the superscript $+$ denotes the pseudoinverse. This solution

backprojects and then filters, while most practical implementations reverse this order. The operators T and T^* can be expanded as sets of orthonormal basis functions, allowing the solution to be rearranged to obtain $f = T^*(TT^*)^+s$ [62]. This is the desired outcome, applying filtering before backprojection. For the forward model given by (2.1) and noiseless data s , the term $(TT^*)^+$ reduces to $(4\pi R)^2$, providing the necessary compensation for the spreading associated with T and T^* .

The preceding result matches backprojection as commonly performed for SAS, SAR, and also basic seismic reconstruction. Furthermore, the well-known Stolt, or ω - k , algorithm can be derived from (2.1) [60]. An important branch of reconstructions can be developed when the imaged scene is interrogated by plane waves or parallel rays. This provides the desired connection to X-ray CT, radio astronomy, and spotlight mode SAR.

For spotlight SAR, the far-field assumption means that the signal reflected from a given range is proportional to the ground reflectivity integrated over a plane, at that range, normal to the radar's line of sight [3]. The recorded signal is called a range profile. The X-ray CT problem is one of transmission, not reflection, and scanners operate on the principle that detectors measure the amount of radiation passing through the patient along straight-line rays from the source. The set of measurements provided at a single time by the detector array is called a projection, or view. According to the projection slice theorem, the 1D Fourier transform of a spotlight range profile or a CT view represents a single slice through the 2D spectrum of the desired image. The connection to radio astronomy is made at this point, because radio interferometry arrays directly measure the 2D Fourier transform of the intensity of the portion of sky being imaged. Reconstruction techniques for all three imaging modalities involve algorithms for taking the Fourier transform of irregular or non-equispaced sampled spectra. This often means interpolating the frequency-domain samples onto a regular sampling grid, suitable for applying the FFT.

CHAPTER 3

ADJOINT ACTIVE SURFACES FOR LOCALIZATION AND IMAGING

This chapter addresses the problem of localizing and segmenting regions embedded within a surrounding medium by characterizing their boundaries, as opposed to imaging the entirety of the volume. Active surfaces are used to directly reconstruct the shape of the region of interest. The optimal surface is computed iteratively via gradient descent that exploits the sensitivity of an error minimization functional to changes of the evolving active surface. The adjoint model is used to compute the sensitivity, and in this respect the method shares common ground with certain other disciplines such as optimal control.

The active surface technique is developed and illustrated in the framework of wave propagation governed by the scalar Helmholtz equation. Two derivations are presented: The first is concerned with finding voids in an infinite medium. The second permits wavefields to exist inside the unknown surfaces and can be used when the surrounding medium is bounded. The former is good for illustrating the approach to deriving active surfaces, while the latter is more general. Both models are relevant to a variety of problems of practical interest. Potential applications include electromagnetics, acoustics, geophysics, nondestructive testing, and medical imaging. Simulated experiments are presented for both models, and experimental results are described in the next chapter.

Since a primary goal of this dissertation is to present adjoint-based active surface reconstruction to the broader signal processing community, the scalar wave equation is chosen as the governing physical model because of its simplicity and ubiquity. The discussion begins by reviewing the adjoint operator, the Green's function, and their application in the context of inverse problems to determine the sensitivity of

an objective function relative to a desired set of inputs. This general concept is used in many disciplines, sometimes explicitly, sometimes not. The forward problem of wave scattering by multiple objects embedded within a homogeneous medium is then introduced, followed by the associated inverse problem formulated by deriving the shape gradient using variational calculus to determine the geometry that minimizes the mismatch between the measured and predicted scattering.

The technique is demonstrated using numerical simulations of circular and spherical regions embedded within a homogeneous medium. Using these shapes permits the use of simple closed-form solutions for the forward and adjoint models based on harmonic expansions. The procedure for dealing with more complicated geometries is the same, although numerical methods such as the boundary element method [63], finite element method [64], or the method of fundamental solutions [65] would be necessary for computing the shape gradient.

3.1 The Adjoint Operator

The adjoint operator is well known in several fields of study as being useful for determining the sensitivity of a measured quantity to changes in some aspect of the governing physics. Notable examples are atmospheric and ocean circulation, optimal control, and geophysics [66, 67, 68, 69]. Fields dominated by traditional Fourier-based signal processing ideas, such as radar and sonar, do not seem to widely exploit the full potential of adjoint analysis for solving their problems. The reason for this is uncertain, although one explanation may be that the traditional approaches, often based on linear time-invariant systems, have worked well for the radar and sonar applications that have been of primary interest since World War II. The adjoint is sometimes found in signal processing applications, but not always explicitly. For example, synthetic aperture backprojection can be formulated in an adjoint context [70, 62] as described in Chapter 2. Looking ahead, emerging problems are proving

too difficult for conventional methods.

Discussions of adjoint operators in the literature take a number of viewpoints. Most of these are somewhat opaque to the reader who is less interested in operator theory and instead simply wants to know what is the practical use of the adjoint. Marchuk [71] and Estep [72] develop the adjoint problem in a rather approachable way: Starting with a differential operator \mathcal{L} acting on a function u , form the inner product by multiplying by an arbitrary test function v and integrating over the domain of interest. Next, use integration by parts to move all derivatives off of u and onto v . The resulting operator on v is the adjoint, which is denoted as \mathcal{L}^* . Instead of motivating this procedure at the outset, the discussion below simply follows it and then explores its usefulness.

Consider a linear differential operator \mathcal{L} operating on a scalar field $u(x, t)$, defined over a spatial domain Ω , with source term $f(x, t)$:

$$\mathcal{L}u = f. \tag{3.1}$$

This operator is accompanied by given boundary conditions $\mathcal{B}u = 0$ on the boundary of Ω , denoted as $\partial\Omega$. Possible boundary conditions include prescribing the value of u , its derivative, or a weighted combination of the two. These are known classically as Dirichlet, Neumann, and Robin boundary conditions, respectively.

The scalar wave equation provides a concrete example, for which $\mathcal{L} = \nabla^2 - c^{-2}\partial/\partial t^2$ in the time domain. The temporal Fourier transform gives the Helmholtz operator $\mathcal{L} = \nabla^2 + k^2$ which is the wave equation for a single frequency ω , where ω is embedded in the wavenumber $k = \omega/c$. The differential operator is always accompanied by appropriate boundary conditions. For the wave equation in an infinite domain Ω , the Sommerfeld radiation condition [73, 74] is used in conjunction with whatever local boundary conditions occur at any scattering surfaces.

The inner product is now formed between (3.1) and a test function v , which is allowed to be arbitrary at this point. The inner product is denoted using angled

brackets:

$$\langle \mathcal{L}u, v \rangle = \int_{\Omega} v \mathcal{L}u \, dx, \quad (3.2)$$

where the integral is taken over the domain Ω of u and v . Integrating by parts as many times as needed (twice for the Helmholtz operator) and applying the given boundary conditions to shift all the derivatives from u and onto v provides a new operator on v [71, 18]. This operator is the adjoint (that is, the operator adjoint to \mathcal{L}) and is denoted by \mathcal{L}^* , together with the associated boundary conditions $\mathcal{B}^*v = 0$. The function v is often also called the adjoint, with the distinction between the adjoint operator \mathcal{L}^* and the adjoint function v usually being clear from context.

The process of integrating by parts results in definite integrals evaluated at the boundary of the domain Ω . Where necessary, the values of v are chosen such that these terms are forced to be zero, leaving the following equality

$$\langle \mathcal{L}u, v \rangle = \langle u, \mathcal{L}^*v \rangle, \quad (3.3)$$

which is known as Lagrange's identity [71]. The requirements imposed on v to obtain (3.3) give the boundary conditions $\mathcal{B}^*v = 0$ that must accompany the adjoint operator \mathcal{L}^* . The adjoint problem is permitted to be inhomogeneous, just as is (3.1)

$$\mathcal{L}^*v = g, \quad (3.4)$$

where the presence and meaning of g is explained in the next section.

In summary, there are two related problems: the main (3.1) and its adjoint (3.4). The main problem is the original physical phenomenon of interest. The second, adjoint, problem is obtained from the steps required to satisfy Lagrange's identity (3.3). The adjoint problem involves the two functions v and g that are soon to be specified. The meaning assigned to these is based on the desire to obtain either the solution of (3.1) or some other particular insight into this problem. It may occur that $\mathcal{L}^* = \mathcal{L}$ and $\mathcal{B}^* = \mathcal{B}$ for certain applications. This is true when (3.1) is a general Helmholtz-type differential operator. Such operators \mathcal{L} are said to be self-adjoint [18].

3.2 The Green's Function

Suppose that the field u may be measured using a process that can be represented as the inner product of u and some measurement function. If so, then this measurement function will be identified with g in (3.4), allowing the properties of the adjoint operator to be exploited. The desired measurement functional is denoted as $J_g = \langle u, g \rangle$, where the subscript indicates that g is the process being used to measure u . Observe that J_g equals the right-hand side of (3.3). In general applications the function g represents the characteristics of an instrument or experiment used to measure u . From (3.1), (3.3), and (3.4) it can be seen that the same measurement may be represented using any of several expressions:

$$J_g = \langle u, g \rangle = \langle u, \mathcal{L}^*v \rangle = \langle \mathcal{L}u, v \rangle = \langle f, v \rangle. \quad (3.5)$$

The best choice among these options depends on the particular objectives and the available data [71].

The present concern is the special case for which u is measured at exactly one point, x_0 . This leads to the Dirac delta function as the choice for g : $g = \delta(x - x_0)$. Acknowledging, but skipping over, the mathematical importance of treating the delta function as a limit of a distribution (3.5) becomes:

$$\begin{aligned} J_g(x_0) &= \langle u(x), \delta(x - x_0) \rangle \\ &= u(x_0). \end{aligned} \quad (3.6)$$

As desired, (3.4) becomes the special relationship

$$\mathcal{L}^*v(x) = \delta(x - x_0), \quad (3.7)$$

with $\mathcal{B}^*v = 0$ on $\partial\Omega$, whose solution v is known as the Green's function and is given the special symbol $G(x, x_0)$.

Assuming that the function $G(x, x_0)$ is known that is transformed into a delta function when operated upon by \mathcal{L}^* , the original problem (3.1) can be solved using

(3.5). That is, the value of the field u at any desired location x_0 can be evaluated by computing the inner product of the source term f and the Green's function:

$$\begin{aligned} u(x_0) &= \langle f(x), G(x, x_0) \rangle \\ &= \int_{\Omega} f(x)G(x, x_0) dx. \end{aligned} \tag{3.8}$$

Because it weights the contribution of the source term $f(x)$ as it is perceived at the point x_0 , the Green's function is sometimes called the influence function. It is also known as the impulse response because u is equal to $G(x, x_0)$ when the source term f is taken to be a delta function (or impulse). The motivation for this terminology becomes even clearer when self-adjoint operators are being studied.

The integral (3.8) is used in electromagnetics where $f(x)$ represents the current distribution on an antenna and $u(x)$ represents the radiated electric field. In the field of acoustics it is known as the Rayleigh integral and is used to predict the sound created by a vibrating source where $f(x)$ represents the velocity of the moving surface [75]. The next section outlines the use of adjoint operators and the associated Green's functions in the solution of inverse problems.

3.3 Adjoint Methods for Inverse Problems

The theory above is now applied to a toy problem to introduce the main idea of the adjoint method. Constraints will be introduced later. For the rigorous formulation of the theory see Marchuk, [71] and [76]. Consider the boundary value problem for u defined on a domain Ω

$$\mathcal{L}_p u = S, \tag{3.9}$$

with Dirichlet boundary conditions, $u = 0$ on $\partial\Omega$. Assume that the differential operator $\mathcal{L}_p = \nabla^2 + k^2(p)$ is of the Helmholtz type and the wavenumber k is function of the unknown parameter vector $p = [p_1, p_2, \dots, p_N]$ that is to be estimated from M

measurements \hat{u}_m at the locations x_m , $m = 1, \dots, M$. To do so, the energy functional

$$E(u, \hat{u}) = \sum_{m=1}^M \frac{1}{2} (u(x_m) - \hat{u}_m)^2 \quad (3.10)$$

representing the mismatch between measurements and the predicted values $u_m = u(x_m)$ via (3.9), can be minimized with respect to the vector of parameters p . The factor of 1/2 is included for convenience, as it cancels the factor of 2 that comes from differentiating the energy.

An artificial time dependence $p(t)$ is introduced to denote the convergence of the parameter estimation. Since the Helmholtz equation is expressed in the frequency domain, this additional usage of the variable t should introduce no confusion. An evolution equation for p can be derived as follows. The time variation of the energy E is given by

$$\frac{dE}{dt} = \sum_{m=1}^M (u(x_m) - \hat{u}_m) \nabla_p u(x_m) \cdot \frac{dp}{dt}, \quad (3.11)$$

which depends upon the gradient $\nabla_p u(x_m)$, where $\nabla_p \equiv [\partial/\partial p_1 \ \partial/\partial p_2 \ \dots \ \partial/\partial p_N]$. The term $\nabla_p u(x_m)$ is the sensitivity of the function u at the location $x = x_m$ due to variations δp in the parameter vector p . It is seen from (3.11) that, in order to reduce E by the maximum possible amount, the parameter p should change in time as

$$\frac{dp}{dt} = -(u(x_m) - \hat{u}_m) \nabla_p u(x_m) \quad (3.12)$$

for each of the M measurements. The key element in such equations is the gradient $\nabla_p u(x_m)$, which can be computed as follows. If the n^{th} parameter p_n changes as $p_n + \delta p_n$ then u varies as $u + \delta u$ according to (3.9) because the solution $u + \delta u$ satisfies (3.9) with p_n replaced by $p_n + \delta p_n$. After neglecting higher order terms since δp_n is infinitesimal, δu satisfies

$$\mathcal{L}_p \delta u = -2k \frac{\partial k}{\partial p_n} \delta p_n u \quad n = 1 \dots N, \quad (3.13)$$

with the condition $\delta u = 0$ imposed on the boundary $\partial\Omega$. Dividing both sides by δp_n

and taking the limit as $\delta p_n \rightarrow 0$ gives

$$\mathcal{L}_p \frac{\partial u}{\partial p_n} = -2k \frac{\partial k}{\partial p_n} u \quad n = 1 \dots N, \quad (3.14)$$

with the boundary condition that $\partial u / \partial p_n = 0$ on the boundary $\partial\Omega$. The solution of this boundary value problem gives the variation in u due to the variation in the n^{th} parameter. The limitation of this approach is that (3.14) has to be solved for each of the N parameters. Thus, in order to compute the MN sensitivities $\nabla_p u(x_m)$, (3.14) needs to be solved N times. This is computationally efficient for well-posed problems since $N \ll M$. However, if $N \gg M$ the computation of these sensitivities becomes slow and cumbersome via (3.14). This is typical of ill-posed inverse problems where the number of parameters N to be estimated is much larger than the number M of available measurements.

The adjoint method accelerates the computation of the sensitivities for the case of $N \gg M$ [77, 78, 67]. The basic idea behind this method is the concept of the Green's function $G(x, x_m)$ of the differential operator \mathcal{L}_p , introduced in the previous section, which satisfies $\mathcal{L}_p^* G = \delta(x - x_m)$ and $G = 0$ on $\partial\Omega$. The function $G(x, x_m)$ can be interpreted as the response of the physical system described by (3.9) to a Dirac delta function located at $x = x_m$. Thus, (3.13) can be solved in terms of $G(x, x_m)$ as

$$\nabla_p u(x_m) = -2k \nabla_p k \int_{\Omega} G(x, x_m) u(x) d\Omega. \quad (3.15)$$

This is the key equation in any adjoint formulation that speeds up the computation of the gradient $\nabla_p u(x_m)$ or, equivalently, the variations $\delta u(x_m)$. For ill-posed problems ($N \gg M$) the adjoint method is more efficient than the standard method based on the direct solution of the sensitivities through (3.14). Indeed, the adjoint technique requires only M solutions of the boundary value problem (3.9) in order to obtain the Green's function for different adjoint source locations x_m . The computation of u then reduces to a simple integration any time one chooses the $N \times 1$ vector of parameters p .

The following discussion introduces the problem of wave scattering by multiple objects within a homogeneous medium (forward problem) and the associated inverse problem of finding the shape of the heterogeneities given measurements of the scattered field. An evolution equation for the object boundary depending on the parameter vector p in (3.12) is derived from the shape gradient. This is the sensitivity of the wave field to changes in the shape of the objects, which is derived using a Lagrange multiplier approach to enforce a differential constraint, similar to (3.9), leading naturally to the adjoint equations.

3.4 Shape Gradient for Opaque Scatterers in an Unbounded Domain

Consider, as shown in Figure 3.1, the unbounded domain Ω partitioned into an outer domain Ω^0 and some number of subdomains Ω^j . The notation for the domains is $\Omega = \Omega^0 \cup \Omega^j$, where the internal boundaries of the j subdomains are denoted Γ^j and the outer boundary of Ω^0 as Γ . Thus, the outer domain (the one containing the subdomains, or unknown objects) is denoted as Ω^0 , its outer boundary is Γ , and its complete boundary is $\Gamma^0 = \Gamma \cup \Gamma^j$. In this example, the domain is unbounded, so the boundary Γ can be thought of as being infinitely far away. Wave propagation within Ω^0 is governed by the Helmholtz equation, and each inclusion Ω^j is a void, on whose surface the field variable u is assumed to be zero. Another common boundary condition is $\partial u / \partial \mathcal{N} = 0$. The domain is unbounded, so u is also taken to be zero on an imaginary surface Γ located very far away. The operator and boundary conditions are summarized as:

$$\mathcal{L}u = S, \quad x \in \Omega^0 \tag{3.16a}$$

$$u = 0, \quad x \in \Gamma^0, \tag{3.16b}$$

where $\mathcal{L}u = \nabla^2 u + k^2 u$.

Since measurements are generally subject to noise, the problem is framed in the

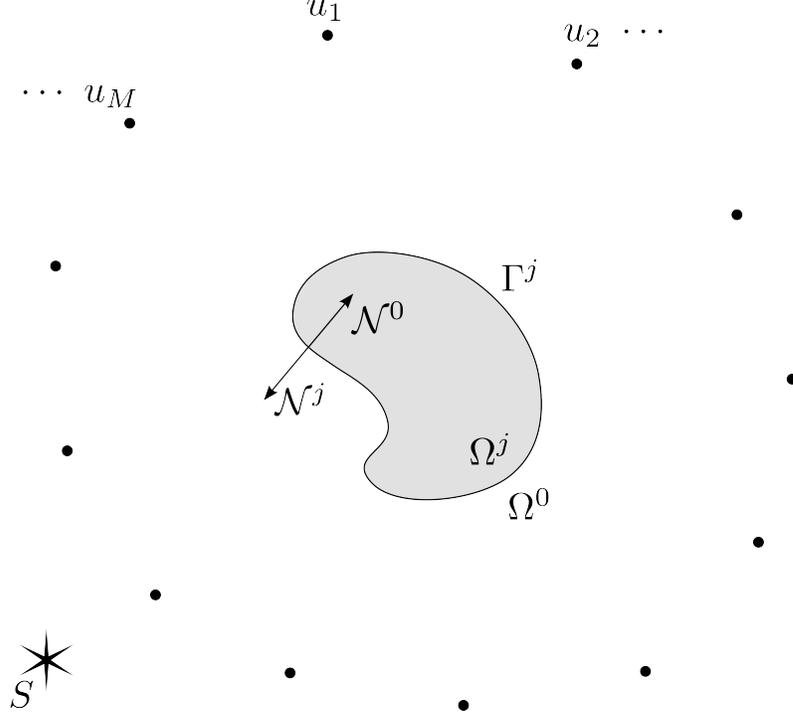


Figure 3.1: Source illuminating an inclusion embedded in a domain Ω^0 containing several measurement locations. Only a single source and inclusion are shown for clarity.

least squares sense by choosing the surface(s) Γ^j to minimize the following energy term representing the summed squared error between the measurements \hat{u}_m and the predicted value u_m at each of the points x_m :

$$E_M = \sum_{m=1}^M \frac{1}{2} (u(x_m) - \hat{u}_m)^2. \quad (3.17)$$

Furthermore, the forward model $\mathcal{L}u = S$ is satisfied exactly as a strong constraint by defining the new energy

$$\begin{aligned} \bar{E}_M &= \sum_{m=1}^M \frac{1}{2} (u(x_m) - \hat{u}_m)^2 + \int_{\Omega} v(\mathcal{L}u - S) dx \\ &= \sum_{m=1}^M \frac{1}{2} (u(x_m) - \hat{u}_m)^2 + \int_{\Omega} v(\nabla^2 u + k^2 u - S) dx \end{aligned} \quad (3.18)$$

The following derivation uses this energy functional to develop a gradient descent algorithm to find the desired Γ^j .

To begin, integration by parts is used to remove the higher-order derivatives of u in (3.18):

$$\bar{E}_M = \sum_{m=1}^M \frac{1}{2} (u(x_m) - \hat{u}_m)^2 + \sum_{j=1}^J \int_{\Gamma^j} v \nabla u \cdot \mathcal{N}^0 ds + \int_{\Omega^0} (-\nabla v \cdot \nabla u + k^2 v u - S v) dx. \quad (3.19)$$

The integration by parts results in integrals over the surfaces bounding Ω^0 . The outer boundary is assumed to be far away, such that the Sommerfeld radiation condition can be used to treat its contribution as zero. This leaves the integrals over the interior surfaces as shown above. These surface integrals need to be eliminated by an appropriate choice of boundary conditions on either u or v . The condition $u = 0$ has been previously specified by the governing physics, so $v = 0$ on Γ^j is the obvious choice. This type of boundary condition corresponds, for example, to acoustic waves interacting with a free surface [79]. (In certain other applications, such as acoustic reflection from rigid boundaries, the normal derivative of u would equal zero, providing a different condition that could remove this term.) Imposing $v = 0$ on the interior boundaries Γ^j leaves

$$\bar{E}_M = \sum_{m=1}^M \frac{1}{2} (u_m - \hat{u}_m)^2 + \int_{\Omega^0} (-\nabla v \cdot \nabla u + k^2 v u - S v) dx. \quad (3.20)$$

The next step is to introduce the artificial time parameter, as in (3.11), and take the derivative with respect to it

$$\begin{aligned} \frac{d\bar{E}_M}{dt} &= \sum_{m=1}^M (u_m - \hat{u}_m) u_{t,m} + \sum_{j=1}^J \int_{\Gamma^j} (-\nabla v \cdot \nabla u + k^2 v u - S v) (\Gamma_t \cdot \mathcal{N}^0) ds \\ &+ \int_{\Omega^0} (-\nabla v_t \cdot \nabla u - \nabla v \cdot \nabla u_t + k^2 v_t u + k^2 v u_t - S v_t) dx. \end{aligned} \quad (3.21)$$

The surface integral above can be simplified by reusing the condition $v = 0$ on Γ^j and noting that the same is true for the source term, so long as the evolving surface is not allowed to reach the location of the source. Making this adjustment and integrating

by parts within the volume integral to remove the gradients from v_t and u_t gives:

$$\begin{aligned} \frac{d\bar{E}_M}{dt} &= \sum_{m=1}^M (u_m - \hat{u}_m) u_{t,m} + \sum_{j=1}^J \int_{\Gamma^j} (-\nabla v \cdot \nabla u)(\Gamma_t \cdot \mathcal{N}^0) ds \\ &\quad + \int_{\Omega^0} (k^2 v_t u + k^2 v u_t - S v_t + v_t \nabla^2 u + u_t \nabla^2 v) dx \\ &\quad - \sum_{j=1}^J \int_{\Gamma^j} (v_t \nabla u + u_t \nabla v) \cdot \mathcal{N}^0 ds \end{aligned} \quad (3.22)$$

The volume integral and summation over the measurement error terms can be eliminated by recalling that $\nabla^2 u + k^2 u - S = 0$ and by imposing as the governing equation for the adjoint, $\nabla^2 v + k^2 v + \sum_{m=1}^M (u_m - \hat{u}_m) = 0$. One of the newly-introduced surface integral terms can also be eliminated reusing $v = 0$ on Γ^j . The result is:

$$\frac{d\bar{E}_M}{dt} = \sum_{j=1}^J \int_{\Gamma^j} (-\nabla v \cdot \nabla u)(\Gamma_t \cdot \mathcal{N}^0) ds - \sum_{j=1}^J \int_{\Gamma^j} (u_t \nabla v) \cdot \mathcal{N}^0 ds. \quad (3.23)$$

The second term may be eliminated since $u = 0$, and therefore $u_t = 0$, on the surface Γ^j . The unit normal vector \mathcal{N}^0 points outward from Ω^0 , so it is convenient to replace it with $-\mathcal{N}^j$, giving

$$\frac{d\bar{E}_M}{dt} = \sum_{j=1}^J \int_{\Gamma^j} (\nabla v \cdot \nabla u)(\Gamma_t \cdot \mathcal{N}^j) ds. \quad (3.24)$$

The derivative $d\bar{E}_M/dt$ is maximized when $\Gamma_t \cdot \mathcal{N}^j = \nabla v \cdot \nabla u$. Since the goal is to minimize this quantity, the desired shape gradient is

$$\Gamma_t = -(\nabla v \cdot \nabla u) \mathcal{N}^j. \quad (3.25)$$

Recall that one of the first steps in the derivation is to choose conditions to eliminate terms that do not contribute to the shape gradient. These conditions are either (1) dictated by the governing physics of the forward model, or (2) imposed on the adjoint function. Whichever approach is taken, a complementary condition will be required later in the derivation. These options are summarized in Table 3.1 for the problem above. It can be seen that a certain symmetry arises from the mathematics in that the physics governing the forward problem are mirrored in the adjoint problem.

Table 3.1: Choices for boundary conditions on the interior surfaces Γ^j used to obtain the shape gradient.

Condition to chosen eliminate surface integral in Equation (3.19)	Condition required to eliminate surface integral in Equation (3.23)
$v = 0$	$u = 0$, or constant
$\frac{\partial u}{\partial \mathcal{N}} = 0$	$\frac{\partial v}{\partial \mathcal{N}} = 0$

3.5 Shape Gradient for Complex Scatterers in a Bounded Domain

This section expands on the result above by (1) deriving the shape gradient for a bounded domain, (2) allowing wave propagation within the scattering regions Ω^j , and (3) generalizing the Helmholtz operator to allow for diffusive media. The approach to deriving the shape gradient is the same, but the algebra is considerably more complicated.

The problem shown in Figure 3.2 is similar to the one shown in Figure 3.1, except that the domain Ω is bounded on its exterior in addition to being partitioned into subdomains Ω^j . As before, the notation for the domains is $\Omega = \Omega^0 \cup \Omega^j$, where the internal boundaries of the j subdomains are denoted Γ^j and the outer boundary as Γ . Thus, the outer domain is denoted as Ω^0 , its outer boundary is Γ , and its complete boundary is $\Gamma^0 = \Gamma \cup \Gamma^j$.

Within the outer domain and the j^{th} subdomain, the material properties are homogeneous and characterized by the pairs (d^0, μ^0) and (d^j, μ^j) respectively. The variables d and μ represent diffusion and propagation constants for the Helmholtz equation. The propagation constant was previously represented as k^2 , but adhering to this more traditional notation would introduce unwanted superscripts in the present context, so $\mu = k^2$ is used instead. The present application is restricted to finding the unknown boundaries, but the material properties could be incorporated

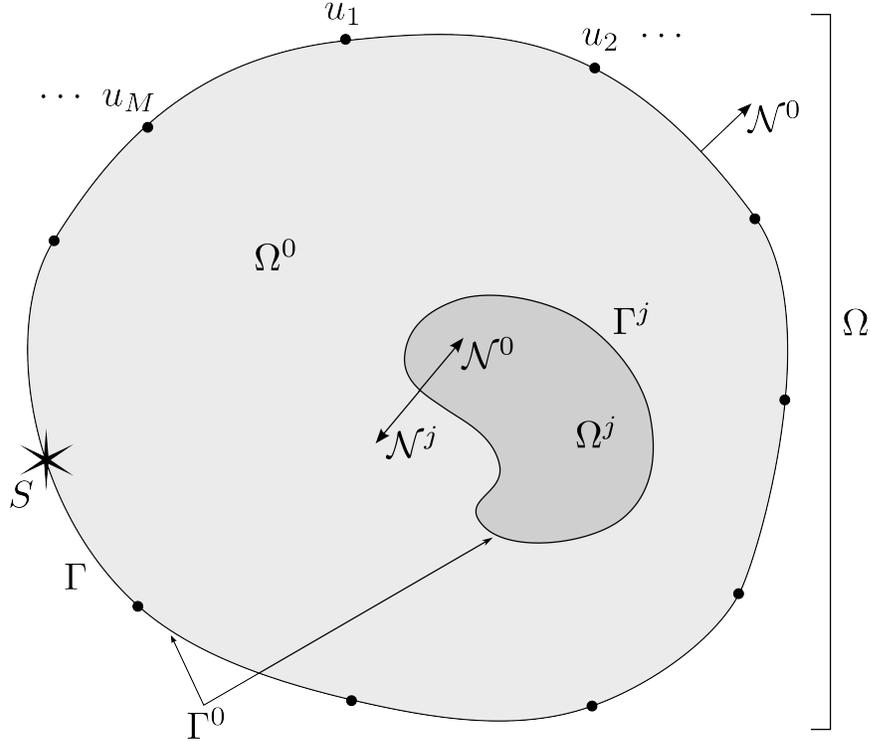


Figure 3.2: Source illuminating an inclusion embedded in a (possibly bounded) domain Ω^0 containing several measurement locations. Only a single source and inclusion are shown for clarity.

into the solution by including them in the vector of unknowns p used previously in the discussion of adjoint solution methods.

The information of ultimate interest is an estimate of the boundaries Γ^j , which could represent the interior structure of a building, objects buried in the sea floor or the Earth, or even tumors within the body. The objective of finding the boundaries is reached somewhat indirectly by meeting a secondary goal of reconstructing the function u over the domain Ω given known excitation source(s) S as well as a set of M measurements $\hat{u}_m = u(x_m) + \eta_m$ at different points $x_m \in \Gamma$, $m = 1, \dots, M$, along the boundary. The η_m terms are random samples from a zero-mean measurement noise process η . The function u represents the total incident and scattered field of the type of radiation appropriate to the problem. It is assumed that u is governed by

the following PDE and Robin boundary condition

$$\mathcal{L}u = S, \quad x \in \Omega \quad (3.26a)$$

$$\frac{\partial u}{\partial \mathcal{N}} + Bu = 0, \quad x \in \Gamma. \quad (3.26b)$$

The operator \mathcal{L} depends upon the partitioning Ω^j , $j = 0, \dots, J$ of the domain Ω resulting from the unknown interior surfaces Γ^j , $j = 1, \dots, J$. It is defined as follows

$$(\mathcal{L}u)(x) = (\mathcal{L}^j u^j)(x), \quad x \in \Omega^j \quad (3.27a)$$

with

$$\mathcal{L}^j \doteq -d^j \nabla^2 + \mu^j, \quad (j = 0, \dots, J), \quad (3.27b)$$

where u^j denotes the restriction of u to Ω^j , with the following matching boundary conditions:

$$u^0 = u^j \quad (3.27c)$$

and

$$d^0 \frac{\partial u^0}{\partial \mathcal{N}^0} = -d^j \frac{\partial u^j}{\partial \mathcal{N}^j}, \quad (3.27d)$$

for $x \in \Gamma^j$, ($j = 1, \dots, J$) imposed across the internal contours Γ^j . These indicate the continuity of u and its normal derivative across each of the subdomain boundaries.

The operator (3.27) describes the propagation of single-frequency waves. The wavenumber k may be complex-valued and is determined by the relationship $k = \sqrt{\mu}$. In typical applications where there is no diffusion d equals -1 , and Equation (3.27b) becomes the ordinary Helmholtz equation. If one wishes to account for attenuation due to absorption, then the wavenumber is complex-valued, with the imaginary part representing the absorption term. An excellent discussion of the physical significance of real and complex-valued wavenumbers can be found in Gumerov [80]. In the event that the constant B in (3.26a) equals zero, the Robin boundary condition collapses to a Dirichlet condition. The Robin boundary condition can be used to impose the Sommerfeld radiation condition on the outer domain [73].

For the sake of convenience, let $d = (d^0, d^1, \dots, d^J)$ and $\mu = (\mu^0, \mu^1, \dots, \mu^J)$ denote the full set of d^j and μ^j coefficients respectively. No similar notation will be used for the set of interior boundaries Γ^j since the symbol Γ is already reserved for the exterior boundary of the complete domain Ω . However, the full set of Γ^j s can still be denoted compactly by noting that $\Gamma^1 \cup \dots \cup \Gamma^J = \Gamma^0 \setminus \Gamma$. A condensed summary, therefore, is that the operator \mathcal{L} in the forward model (3.26a) is defined by (3.27) once d , μ , and $\Gamma^0 \setminus \Gamma$ are given.

The inverse problem may now be stated as that of determining the unknown interior partition surfaces $\Gamma^0 \setminus \Gamma$ which produce an operator \mathcal{L} , via (3.27), that yields a function u whose values at each x_m match the known measurements \hat{u}_m once the corresponding forward model (3.26a) has been solved. It is assumed that the coefficients d and μ are known.

As before, the desired matching problem is posed in the least squares sense by choosing the surface $\Gamma^0 \setminus \Gamma$ to minimize the energy term representing the summed squared error between the measurements \hat{u}_m and the true value u at each of the points x_m :

$$E_M = \sum_{m=1}^M \frac{1}{2} (u(x_m) - \hat{u}_m)^2. \quad (3.28)$$

This equation corresponds to (3.10), and the goal is to derive a gradient descent based on the idea of changing the unknown boundary $\Gamma^0 \setminus \Gamma$ until E_M is minimized. The sifting property of the delta function is used to write the energy E_M in integral form as

$$E_M = \int_{\Gamma} g(u(x), x) ds, \quad (3.29)$$

where

$$g(u, x) = \sum_{m=1}^M \frac{1}{2} (u - \hat{u}_m)^2 \delta(x - x_m). \quad (3.30)$$

Furthermore, equation (3.26a) (the forward model $\mathcal{L}u = S$) is satisfied exactly as a

strong constraint by defining the new energy

$$\bar{E}_M = \int_{\Gamma} g ds + \int_{\Omega} v(\mathcal{L}u - S) dx. \quad (3.31)$$

The function v is a Lagrange multiplier and will be shown to be the adjoint of u . The method of constraint (3.31) contrasts with a weaker form in which a penalty function combined with a user-defined constant would be appended to (3.29).

To carry out the variational calculations necessary to obtain the shape gradient, Equation (3.31) is rewritten for \bar{E}_M more directly in terms of the unknown parameters, integration by parts is performed, and then the internal and external boundary integrals are separated resulting in (3.32),

$$\begin{aligned} \bar{E}_M &= \int_{\Gamma} g ds + \sum_{j=0}^J \int_{\Omega^j} v^j (-d^j \nabla^2 u^j + \mu^j u^j - S) dx \\ &= \int_{\Gamma} (g + d^0 B v^0 u^0) ds + \sum_{j=1}^J \int_{\Gamma^j} d^j (v^0 - v^j) \nabla u^j \cdot \mathcal{N}^j ds \\ &\quad + \sum_{j=0}^J \int_{\Omega^j} (d^j \nabla v^j \cdot \nabla u^j + \mu^j v^j u^j - v^j S) dx. \end{aligned} \quad (3.32)$$

The boundary conditions (3.26b) for u were used to simplify the exterior Γ -integral, and the flux conditions (3.27d) for \mathcal{L} were used to simplify the interior Γ^j -integrals.

The interior boundary integrals can then be eliminated by imposing the same continuity condition (3.27c) for the test function v :

$$v^0 = v^j, \quad x \in \Gamma^j, \quad (3.33)$$

leaving

$$\begin{aligned} \bar{E}_M &= \int_{\Gamma} (g + d^0 B v^0 u^0) ds \\ &\quad + \sum_{j=0}^J \int_{\Omega^j} (d^j \nabla v^j \cdot \nabla u^j + \mu^j v^j u^j - v^j S) dx; \end{aligned} \quad (3.34)$$

a simplified expression for \bar{E}_M which is used to develop the shape gradient. As discussed in Section 3.1, the process of integrating by parts and imposing boundary conditions on v provides the adjoint problem.

Assuming knowledge of d and μ and an initial guess for the unknown surface $\Gamma^0 \setminus \Gamma$, the variation of \bar{E}_M is determined given a perturbation of the j^{th} interior boundary Γ^j . This is accomplished by introducing the artificial time variable t and by letting Γ^j be time dependent. As a result, the operator \mathcal{L} also becomes time dependent and therefore so does u since conditions (3.26a) continue to be imposed, which are determined by Γ . Maintaining this constraint allows the time derivatives of E_M to be equated with \bar{E}_M which is related to the perturbation $\Gamma_t^j = \partial\Gamma^j/\partial t$ of the surface Γ^j (non-integer variables as subscripts denote partial derivatives with respect to the corresponding variable). Finally, the unknown test function v is made to be time dependent, while continuing to impose the continuity constraint (3.33). The time derivative of the matching energy (3.34) can now be computed, which is shown in (3.35),

$$\begin{aligned}
\frac{d\bar{E}_M}{dt} &= \int_{\Gamma} \left(g_u u_t^0 + d^0 B(v_t^0 u^0 + v^0 u_t^0) \right) ds \\
&\quad + \sum_{j=0}^J \int_{\Gamma^j} (d^j \nabla v^j \cdot \nabla u^j + \mu^j v^j u^j - v^j S)(\Gamma_t^j \cdot \mathcal{N}^j) ds \\
&\quad + \sum_{j=0}^J \int_{\Omega^j} \left((d^j \nabla v_t^j \cdot \nabla u^j + \mu^j v_t^j u^j - v_t^j S) + (d^j \nabla v^j \cdot \nabla u_t^j + \mu^j v^j u_t^j) \right) dx.
\end{aligned} \tag{3.35}$$

It is not possible to directly compute the minimizing boundary surfaces. Instead a gradient descent procedure is devised in which the change in \bar{E}_M is related to changes in the surface $\Gamma^0 \setminus \Gamma$. This is accomplished using the artificial time dependence $\bar{E}_M(t)$ and carrying out a series of computations to arrive at an expression for $d\bar{E}_M/dt$. It is important to remember that the time dependence used here is unrelated to any interpretation of time found in the governing physics. Recall that the Helmholtz equation does not depend on time: It represents harmonic oscillation at a single frequency that depends only on the boundary conditions.

The artificial time dependence is imposed on the underlying functions u and v and therefore carries through to Γ^j and \bar{E}_M so that solving $d\bar{E}_M/dt = 0$ means that

the optimal estimate of the boundaries has been found. Each step of the gradient descent requires knowledge of $\partial\Gamma^j/\partial t$, which is termed the shape gradient.

The process of deriving $d\bar{E}_M/dt$ allows the sought-after $\partial\Gamma^j/\partial t$ to be isolated and expressed in terms of the known quantities u and v . These may be computed because each is the solution to a well-posed problem consisting of a governing PDE and appropriate boundary conditions. In general u and v would be found numerically, but the special case of a spherical inhomogeneity has a closed-form solution expressed in terms of spherical harmonics [81].

After first applying integration by parts to the volume integrals in (3.35), the result may be simplified by noting that $\mathcal{L}^j u^j = 0$ and by imposing the following condition on the adjoint function v :

$$\mathcal{L}^j v^j = -d^j \nabla^2 v^j + \mu^j v^j = 0, \quad x \in \Omega^j, \quad (3.36)$$

thereby eliminating the volume integral terms. The surface Γ^0 is then expanded into its constituent terms, permitting the elimination of the exterior boundary integral (along Γ) by noting the Robin boundary conditions (3.26ab) for u and by choosing the following boundary conditions for v ,

$$d^0 \left(Bv + \frac{\partial v}{\partial \mathcal{N}} \right) = -g_u, \quad x \in \Gamma. \quad (3.37)$$

The term g_u is the partial derivative of (3.30) with respect to u . Imposing (3.37), replacing u^0 and v^0 with u^j and v^j respectively (since they must match at the boundary Γ^j), as well as substituting $\mathcal{N}^0 = -\mathcal{N}^j$ yields (3.38) from (3.35),

$$\begin{aligned} \frac{d\bar{E}_M}{dt} &= \sum_{j=1}^J \int_{\Gamma^j} \left((d^j \nabla v^j \cdot \nabla u^j - d^0 \nabla v^0 \cdot \nabla u^0) + (\mu^j - \mu^0) v^j u^j \right) (\Gamma_t^j \cdot \mathcal{N}^j) ds \\ &\quad + \sum_{j=1}^J \int_{\Gamma^j} \left(d^0 \left(v_t^0 \frac{\partial u^0}{\partial \mathcal{N}^0} + \frac{\partial v^0}{\partial \mathcal{N}^0} u_t^0 \right) + d^j \left(v_t^j \frac{\partial u^j}{\partial \mathcal{N}^j} + \frac{\partial v^j}{\partial \mathcal{N}^j} u_t^j \right) \right) ds. \end{aligned} \quad (3.38)$$

Continuing the calculation requires the relationship between u_t^0 and u_t^j along the moving boundary $\Gamma^j(t)$, which is obtained by differentiating the continuity conditions

(3.27b) for u and (3.33) for v , as shown in (3.39),

$$\begin{aligned} \frac{d}{dt} \left[u^0 \left(\Gamma^j(t), t \right) = u^j \left(\Gamma^j(t), t \right) \right] &\longrightarrow \\ u_t^0 &= u_t^j + (\nabla u^j - \nabla u^0) \cdot \Gamma_t^j \\ &= u_t^j + \left(\frac{\partial u^j}{\partial \mathcal{N}^j} - \frac{\partial u^0}{\partial \mathcal{N}^j} \right) (\Gamma_t^j \cdot \mathcal{N}^j) \end{aligned} \quad (3.39)$$

Substituting these derivatives into (3.38) gives an expression that can be further simplified by utilizing the flux condition (3.27d) for u and by imposing the same condition as the final constraint on v ,

$$d^0 \frac{\partial v^0}{\partial \mathcal{N}^0} = -d^j \frac{\partial v^j}{\partial \mathcal{N}^j}, \quad x \in \Gamma^j, \quad (3.40)$$

finally obtaining (3.41).

$$\begin{aligned} \frac{d\bar{E}_M}{dt} &= \sum_{j=1}^J \int_{\Gamma^j} \left((d^j \nabla v^j \cdot \nabla u^j - d^0 \nabla v^0 \cdot \nabla u^0) + (\mu^j - \mu^0) v^j u^j \right) (\Gamma_t^j \cdot \mathcal{N}^j) ds \\ &\quad + \sum_{j=1}^J \int_{\Gamma^j} \left(- \left(\frac{\partial v^j}{\partial \mathcal{N}^j} + \frac{\partial v^0}{\partial \mathcal{N}^0} \right) d^j \frac{\partial u^j}{\partial \mathcal{N}^j} + \left(\frac{\partial u^j}{\partial \mathcal{N}^j} + \frac{\partial u^0}{\partial \mathcal{N}^0} \right) d^0 \frac{\partial v^0}{\partial \mathcal{N}^0} \right) (\Gamma_t^j \cdot \mathcal{N}^j) ds \end{aligned} \quad (3.41)$$

The shape sensitivity calculation is now completed by recognizing that (3.41) is in the form of a directional derivative: It provides the the rate of change of \bar{E}_M in terms of the rate of change of the surface Γ^j . The quantity $d\bar{E}_M/dt$ is maximized if Γ_t^j is set equal to the remaining terms on the right-hand side of (3.41). The energy is minimized by moving Γ^j in the opposite direction. The the shape gradient for an arbitrary 3D surface boundary Γ^j is therefore given by (3.42).

$$\begin{aligned} \frac{\partial \Gamma^j}{\partial t} &= \left[\left(\frac{\partial v^j}{\partial \mathcal{N}^j} + \frac{\partial v^0}{\partial \mathcal{N}^0} \right) d^j \frac{\partial u^j}{\partial \mathcal{N}^j} - \left(\frac{\partial u^j}{\partial \mathcal{N}^j} + \frac{\partial u^0}{\partial \mathcal{N}^0} \right) d^0 \frac{\partial v^0}{\partial \mathcal{N}^0} \right. \\ &\quad \left. + (d^0 \nabla v^0 \cdot \nabla u^0 - d^j \nabla v^j \cdot \nabla u^j) + (\mu^0 - \mu^j) v^j u^j \right] \mathcal{N}^j. \end{aligned} \quad (3.42)$$

The constraints (3.33) and (3.40) on v along the internal contours Γ^j , $j = 1, \dots, J$ are identical to the constraints associated with the forward model operator \mathcal{L} , and

the local operators \mathcal{L}^j applied to v^j over each region Ω^j in the PDE constraint (3.36) are also identical to those used by \mathcal{L} . As a result, the overall composite operator used on v is the same composite operator \mathcal{L} defined in (3.27) for u . Thus, together with the boundary conditions (3.37), the constraints (3.33), (3.36), and (3.40) constitute a well-posed problem for v , allowing a unique solution to be computed. The model for v may be stated compactly as

$$\begin{aligned} \mathcal{L}v &= 0, \quad x \in \Omega \\ d^0 \left(\frac{\partial v}{\partial \mathcal{N}} + Bv \right) + \sum_{m=1}^M (u - \hat{u}_m) \delta(x - x_m) &= 0, \quad x \in \Gamma. \end{aligned} \quad (3.43)$$

The Lagrange multiplier v is the adjoint of u which arises naturally from the Lagrangian constrained minimization of (3.31). Further, it is seen that the operator for v is the same as that of the forward problem. Therefore, \mathcal{L} and its boundary conditions constitute a self-adjoint operator.

The terms involving normal derivatives within the square brackets of (3.42) can be simplified further. First substitute $-\mathcal{N}^j$ for the inward normal \mathcal{N}^0 , and then use the fact that the normal derivative is equal to the inner product of the gradient and the unit normal vector. For example,

$$\left(\frac{\partial v^j}{\partial \mathcal{N}^j} - \frac{\partial v^0}{\partial \mathcal{N}^j} \right) \frac{\partial u^j}{\partial \mathcal{N}^j} = \left(\frac{\partial v^j}{\partial \mathcal{N}^j} - \frac{\partial v^0}{\partial \mathcal{N}^j} \right) \mathcal{N}^j \cdot \nabla u^j. \quad (3.44)$$

The terms enclosed by parentheses contain only normal derivatives. However, the boundary conditions require that $v^j = v^0$, so it follows that the components of the gradient in the plane tangent to the surface Γ^j are equal as well. This allows (3.44) to be rewritten as

$$\begin{aligned} &\left(\frac{\partial v^j}{\partial \mathcal{N}^j} - \frac{\partial v^0}{\partial \mathcal{N}^j} \right) \mathcal{N}^j \cdot \nabla u^j \\ &= \left[\left(\frac{\partial v^j}{\partial \mathcal{N}^j} - \frac{\partial v^0}{\partial \mathcal{N}^j} \right) \mathcal{N}^j + \left(\frac{\partial v^j}{\partial \mathcal{T}^j} - \frac{\partial v^0}{\partial \mathcal{T}^j} \right) \mathcal{T}^j \right] \cdot \nabla u^j \\ &= (\nabla v^j - \nabla v^0) \cdot \nabla u^j. \end{aligned} \quad (3.45)$$

The additional terms expressing the components of the gradient in the plane tangent to Γ^j sum to zero, but they allow for a more compact notation. The same argument can be applied to the other term in parentheses in (3.42) to arrive at the following form of the shape gradient:

$$\frac{\partial \Gamma^j}{\partial t} = \left[(d^0 - d^j) \nabla u^j \cdot \nabla v^0 + (\mu^0 - \mu^j) v^j u^j \right] \mathcal{N}^j. \quad (3.46)$$

The shape gradient is a function of both u and v . The forward model provides knowledge of u , where u is computed using the current estimate of the scattering environment. The mathematics in this section show that v is the solution to the operator and boundary conditions adjoint to those associated with u . The source terms for v are the differences between the measured and predicted fields at the locations x_m , given by (3.30). Because the operators (3.27) and (3.43) are the same, the adjoint field v can be computed using the same method used to solve the forward model. There is no requirement for explicitly inverting the forward model, thus demonstrating one of the significant strengths of adjoint-based approaches.

CHAPTER 4

NUMERICAL EXPERIMENTS FOR ADJOINT ACTIVE SURFACES

The derivation of the shape gradient applies to any surface in the presence of wave propagation governed by the Helmholtz equation. For the purpose of illustration the results of Section 3.5 are applied to the problem of locating the position and radius of a spherical inclusion embedded in free space. This approach should also work well in situations where the object can be approximated as a sphere. This is the case whenever the characteristic dimension a of the target is much smaller than the wavelength used to interrogate the volume ($ka \ll 1$). The material properties of the medium and the inclusion are known. Measurements of single-frequency (i.e., monochromatic) scattered waves are made in a region surrounding the location of the inclusion.

4.1 Closed-Form Solution for Forward and Adjoint Problems

Chapter 16 of Jackson [81] gives the following classical eigenfunction expansion for the solution of the forward (3.27) and adjoint (3.43) problems when concerned with scattering from a sphere:

$$\begin{aligned}
 u^0(r, \phi, \theta) &= G^0(R) + \sum_{n=0}^{\infty} \sum_{m=-n}^n Y_n^m(\phi, \theta) [A_{mn} h_n^{(1)}(k_0 r) + B_{mn} h_n^{(2)}(k_0 r)] \\
 u^j(r, \phi, \theta) &= \sum_{n=0}^{\infty} \sum_{m=-n}^n C_{mn} Y_n^m(\phi, \theta) j_n(k_j r),
 \end{aligned} \tag{4.1}$$

where k_0 and k_j are the usual wavenumbers expressing spatial frequency. The terms $h_n^{(1)}(k_0 r)$ and $h_n^{(2)}(k_0 r)$ are the spherical Hankel functions, and $j_n(k_j r)$ is the spherical Bessel function of the first kind. These are generated by combining the Bessel functions of the first and second kinds as shown in Table 4.1. The symbol G^0 represents the free-space Green's function $\exp(ik^0 R)/4\pi R$, Y_m^n are the spherical harmonics, and A_{mn} , B_{mn} , and C_{mn} are the Fourier coefficients of the expansion which are found

Table 4.1: Spherical Hankel functions defined in terms of ordinary Bessel functions of the first and second kinds.

Bessel function of the first kind:	$J_n(x)$
Bessel function of the second kind:	$N_n(x)$
Spherical Bessel function of the first kind:	$j_n(x) = \frac{1}{\sqrt{x}} J_{n+\frac{1}{2}}(x)$
Spherical Bessel function of the second kind:	$n_n(x) = \frac{1}{\sqrt{x}} N_{n+\frac{1}{2}}(x)$
Spherical Hankel function of the first kind:	$h_n^{(1)}(x) = j_n(x) + in_n(x)$
Spherical Hankel function of the second kind:	$h_n^{(2)}(x) = j_n(x) - in_n(x)$

using the boundary conditions. While spherical harmonic expansions are generally little-used by the signal processing community, they are relatively simple to work with, being analogous to more common Fourier series expansions in terms of sines and cosines. The first few spherical harmonics are shown in Figure 4.1.

The pair of Hankel functions represents outward (+) and inward (−) propagating waves in a way that is directly analogous to the more familiar representation of harmonic free space spherical waves, $G(R) = \exp(\pm ikR)/4\pi R$. The spherical harmonic terms $Y_n^m(\phi, \theta)$ are given by:

$$Y_n^m(\phi, \theta) = \sqrt{\frac{(2m+1)(n-m)!}{4\pi(n+m)!}} e^{im\phi} P_n^m(\theta), \quad (4.2)$$

where $P_n^m(\theta)$ are the Legendre functions of degree n and order m .

The free-space Green's function $G^0(R)$ appearing in (4.1) is a function of R , the distance between the source and the point of observation. The remaining terms in (4.1) depend on r , ϕ and θ , which are the spherical coordinates representing radius from the origin, azimuth angle, and colatitude (or zenith angle). The origin is located at the center of the scattering sphere, and the colatitude is referenced to the line connecting the source and the scattering sphere.

It is useful to eliminate the dependence on R by expanding $G^0(R)$ in terms of these variables using the spherical Bessel functions and Legendre polynomials (i.e.,

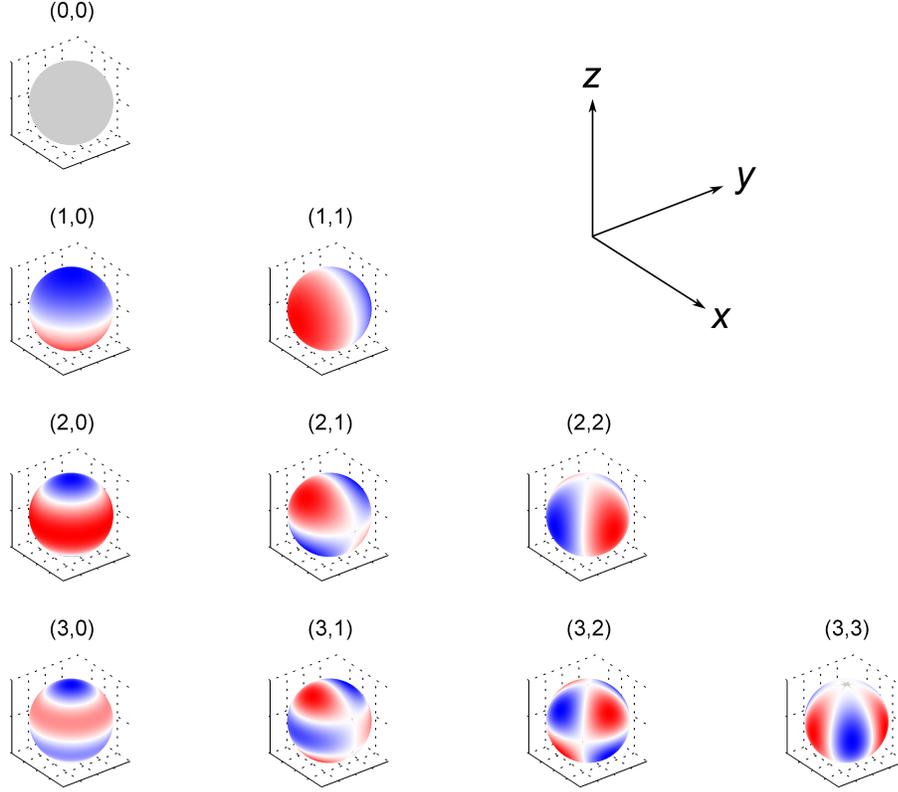


Figure 4.1: Spherical harmonics of degree n and order m . The harmonics $(n, 0)$ are axisymmetric about the z -axis and are known as zonal harmonics, as they have no zero crossings along lines of constant latitude. Those corresponding to (n, n) are called sectoral harmonics and have no zero crossings along lines of constant longitude. The remaining harmonics (n, m) for $m \neq 0$ and $m < n$ have zero crossings in both directions and are called tesseral harmonics [80].

the Legendre functions $P_n^m(\theta)$ for which $m = 0$),

$$\begin{aligned}
 G^0(R) &= \frac{S}{4\pi d^0} \frac{e^{ik_0 R}}{R} \\
 &= \frac{iSk_0}{4\pi d^0} \sum_{n=0}^{\infty} (2n+1) j_n(k_0 r) h_n^{(1)}(k_0 \rho) P_n(\theta).
 \end{aligned} \tag{4.3}$$

This expansion replaces the dependence on R with dependence on r, ρ , and θ , thus matching quantities used in the other terms in (4.1). The variable ρ is new, but known: It represents the distance from the source S to the origin, which is located at the center of the inclusion illustrated in Figure 4.2.

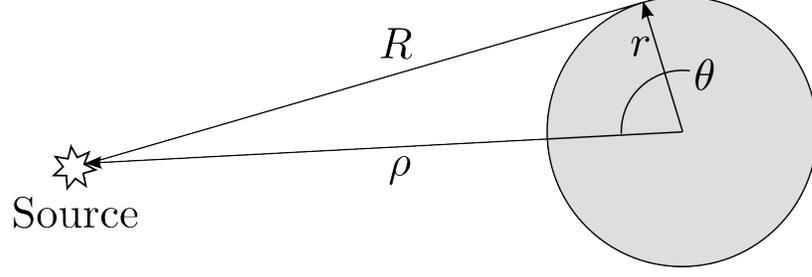


Figure 4.2: Coordinates used for the spherical harmonic expansion of u and v . The distance from the source to a point on the surface of Γ^j is R , while ρ is the distance from center of Γ^j to the source, and θ is the colatitude, or zenith angle. The latitude, or azimuth angle, is not shown but is measured about the axis labeled ρ .

For this example the outer domain is assumed to be free space, so Γ is considered to be infinitely far away from the source S . In this case, the solution u satisfies the Sommerfeld radiation condition. According to this radiation condition, the unbounded outer domain precludes the propagation of any inward-travelling waves. Consequently, the coefficients B_{mn} are all equal to zero since only the outward propagating solution is allowed. Furthermore, the scalar field u and its normal derivative must be continuous at the surface of the spherical inclusion:

$$u^0 = u^j, \quad \text{and} \quad d^0 \frac{\partial u^0}{\partial \mathcal{N}^0} = -d^j \frac{\partial u^j}{\partial \mathcal{N}^j} \quad \text{on } \Gamma. \quad (4.4)$$

Since Γ^j is a sphere, the outward normal derivative is simply the radial derivative $\partial/\partial \mathcal{N}^j = \partial/\partial r$. The resulting configuration is also axisymmetric, causing the dependence on the azimuth angle ϕ to vanish. The axis of symmetry is the line of length ρ connecting the source to the center of the spherical object in Figure 4.2.

Simplifying the expansions above and substituting them into the matching boundary conditions (4.4) gives, for each term in the expansion:

$$\begin{aligned} \frac{ik_0 S}{d^0} \sqrt{\frac{2n+1}{4\pi}} j_n(k_0 r^j) h_n^{(1)}(k_0 \rho) + A_n h_n^{(1)}(k_0 r^j) &= C_n j_n(k_j r^j) \\ ik_0^2 S \sqrt{\frac{2n+1}{4\pi}} j_n'(k_0 r^j) h_n^{(1)}(k_0 \rho) + A_n d^0 k_0 h_n'^{(1)}(k_0 r^j) &= C_n d^j k_j j_n'(k_j r^j). \end{aligned} \quad (4.5)$$

Due to axisymmetry the only nonzero values of the expansion coefficients correspond

to the index $m = 0$ since the exponential function and Legendre function within the spherical harmonics (4.2) are constant for all azimuth angles only when $m = 0$. The m index is suppressed for notational simplicity: For example, A_{0n} is written as A_n in (4.5).

Equations (4.5) can be arranged into a 2×2 system of linear equations which can be easily solved explicitly for the coefficients $[A_n \ C_n]^T$ for a single term in the expansion. Given the source strength and position, along with the position and radius of the sphere Ω^j , the scattered field is given in closed form by summing over the terms in the harmonic expansion.

Computing the value of Γ_t^j for a spherical flaw embedded in free space requires knowledge of the fields u and v both inside and outside the sphere. Equations (4.5) are used to compute u^0 and u^j . They also serve as a prototype for computing the adjoint field v , in which the measurement errors $S_m = u(x_m) - \hat{u}_m(x_m)$ act as source terms replacing S in (4.5) and where the spherical Hankel function of the second kind replaces $h_n^{(1)}(z)$.

The forward and adjoint fields generally have different axes of symmetry since the vectors from the center of Ω^j to the respective sources and measurement locations are different. It is therefore necessary to transform from the local spherical coordinates into a common frame of reference when computing the required scalar fields and their gradients. This is conveniently done using the axis-angle representation for a rotation matrix, where the entries of the rotation matrix are computed directly from the components of the unit vector defining the axis of rotation and the angle through which to rotate.

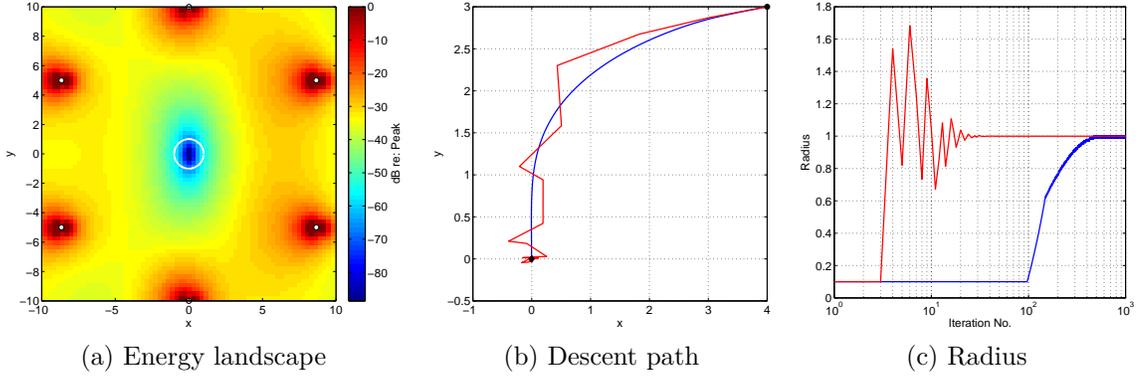


Figure 4.3: Left: The energy landscape constructed by computing the energy \bar{E}_M over a region of the (x, y) plane where $x \in [-10, 10]$ and $y \in [-10, 10]$. The plot is shown in terms of decibels relative to the peak value of \bar{E}_M within the computed region. The measurement locations are indicated by the white markers. Center and Right: Descent path and corresponding radius as the gradient descent algorithm reduces the energy \bar{E}_M . The plots compare results obtained using a fixed step size of 0.01 units (dashed blue curves) with an adaptive scheme (solid red curves) in which a larger step size grows (or shrinks) in response to whether \bar{E}_M decreased (or increased) in the previous step.

4.2 Closed-Form Solution for Spherical Objects

The example shown in Figure 4.3 consists of a single source and a ring of six measurement positions. The source is at a distance of 1,000 units from the ring of measurements, which is centered on the origin with radius of 10 units. Generic units of length are used for the purpose of illustration. The length scales would be normalized by the wavelength to translate these results into a specific application.

The coefficient d is dimensionless, while μ has dimension of inverse length squared. The material properties used for this example are $d^0 = -1$, $d^j = -1 \times 10^9$, $\mu^0 = 0.1$, and $\mu^i = 1$. Since $k = 2\pi/\lambda$ the wavelength for this example is approximately 62.8 units.

For this example the true sphere has unit radius and is located at the origin. Experiments also showed that the method worked well when the true sphere location was offset from the origin. An energy landscape can be visualized by marching the current guess of the sphere's position over a rectangular grid spanning the region of

interest, as in Figure 4.3a. The summed squared difference between the measured and predicted scattered field is plotted for each point. As expected, the minimum energy occurs at the origin. Although it is not depicted, performing this same computation using the incorrect radius can cause the minimum to shift away from the true position of the object. Such effects must be considered when devising the gradient descent scheme to be implemented.

The unknown boundary of the sphere embedded in free space is found iteratively via gradient descent based on (3.46). The current estimate of the boundary at each step is used to model the scattered field, which is compared with the observed measurements. This mismatch provides the source terms used in the adjoint formulation which backpropagates the adjoint field to the surface of the current guess Γ^j . The forward-modeled and adjoint fields are used to find the change in the summed squared error that would result from a given perturbation of the flaw position and radius, as previously described. The general algorithm is as follows:

Make an initial guess for the object boundaries Γ^j . From this, compute the initial values of u and v ;

while Stopping condition is not satisfied **do**

 Compute Γ_t^j based on current values of Γ^j , v , and u ;

 Update Γ^j ;

 Compute updated v and u based on new Γ^j ;

end while

The stopping condition is reached when the norm of Γ_t^j falls below some specified threshold or when a maximum number of iterations have occurred.

Since the object is constrained to be a sphere, the shape gradient Γ_t^j can be expressed in terms of position by integrating $\Gamma_t^j \cdot e_n$ over the sphere, where e_n is the

unit vector in the x , y , or z direction. Similarly, the radial shape gradient is found by integrating $\Gamma_t^j \cdot \mathcal{N}^j$ over the sphere. The sphere was discretized for the numerical computations, so the integrals were replaced by summations. The fact that very long wavelengths can be used to accurately locate the sphere indicates that the use of prior information has a regularizing effect on the solution. Similar ideas are applied by Nain et al. to the use of active contours for segmenting medical imagery in [82].

Figures 4.3b and 4.3c show a gradient descent path for which the initial guess position is $(x, y) = (4, 3)$ units, and for which the radius is initially guessed to be 10 times too small ($r = 0.1$ units). The solution converges, making several pauses along the way as the radius updates. This is an artifact of the specific gradient descent procedure used.

To investigate the effects of noisy data, white Gaussian noise was added to each measurement to achieve average signal to noise ratio (SNR) values of 20, 30 and 40 dB relative to the measured signal power, and the results are shown in Figure 4.4. As in Figure 4.3a, the energy landscapes plotted in this figure used the correct value for the radius of the sphere. The actual sphere location is shown by the solid line, while the dashed line shows the location implied by the minimum of the energy landscape. One concludes from Figure 4.4 that noise can profoundly affect the accuracy of the localization. In an attempt to mitigate the degraded performance, using the 20 dB case as an example, the number of measurement locations was increased from 12 to 24 as shown in Figure 4.5. As expected, adding more measurement locations tends to average out the noise and restore the global minimum located at the origin. In all of these results, only a single measurement was used at each observation point. Time averaging is normally used in practice to improve the SNR at each location. The effects of noise on similar reconstructions are reported by Dorn et al. [51] and He et al. [83].

Another means of mitigating the impact of noise is to repeat the experiment over

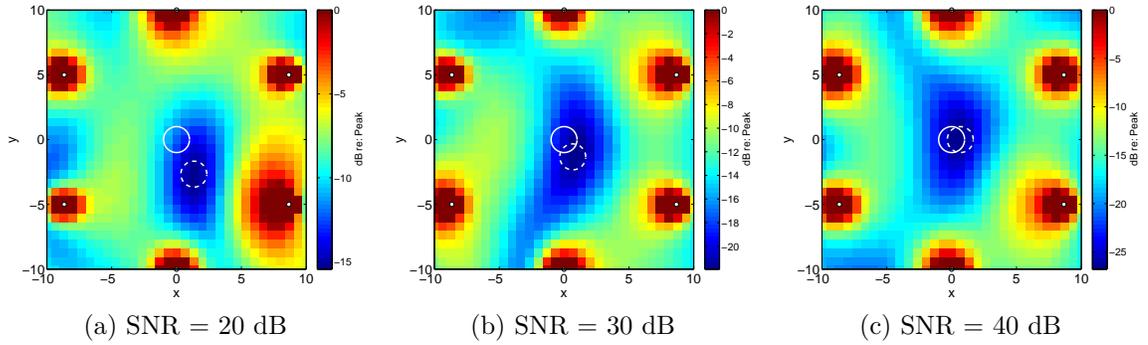


Figure 4.4: Inversion of noisy data: Energy landscapes for SNR = (a) 20 dB, (b) 30 dB, and (c) 40 dB. The true sphere location is shown as a solid line and the location implied by the minimum of the energy landscape as a dashed line. Note that the dynamic ranges are significantly smaller than that used in Figure 4.3a.

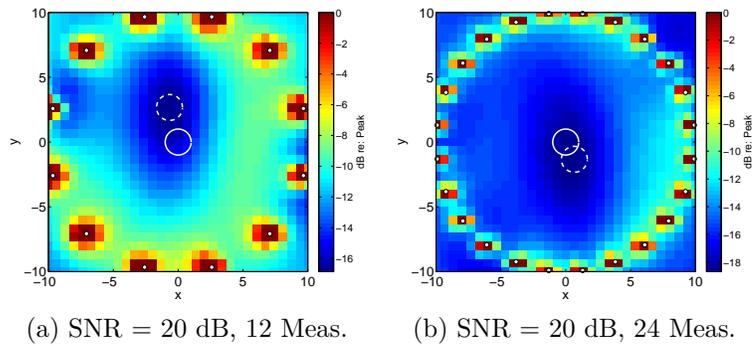


Figure 4.5: Energy landscapes for noisy data (SNR=20 dB) with (a) 12 and (b) 24 measurements. The actual sphere location is shown as a solid line, and the location implied by the minimum of the energy landscape as a dashed line.

a band of frequencies. In this case, one would solve the forward and adjoint problem at each of a discrete set of frequencies and then sum the results. This approach is also how one would handle a transient signal since the temporal Fourier transform provides the contribution of each frequency present. One interesting approach found in the literature [84, 85] begins with a long-wavelength radiation for the initial reconstruction and then moves to increasingly shorter wavelengths to refine the details. This ‘frequency hopping’ approach serves as a kind of time-dependent regularization for the reconstruction that relaxes as the solution converges. Such frequency diversity could be used either in lieu of or in conjunction with any prior information regarding the target domain.

Reassuringly, the variational adjoint method appears to follow the same sampling principles as traditional array processing in that the wavefield needs to be sampled at intervals equal to or smaller than one-half of the wavelength. Local minima begin to appear if the distance between measurement locations is increased. These are presumably analogous to the ambiguities that occur in a traditional array when it is spatially undersampled.

The preceding discussion employed isotropic sources and measurements. If directivity, or beam pattern, effects were considered these would limit the spatial frequencies visible to the sensors and the measurements could perhaps be separated by more than $\lambda/2$ [26]. These observations are based on numerical experiments. The precise relationship between variational techniques and traditional sampling requirements remains to be explored in the literature.

CHAPTER 5

ADJOINT ACTIVE SURFACES FOR SCATTERING OF ULTRASONIC LAMB WAVES

This chapter applies the principles of adjoint active surface imaging to the scattering of ultrasonic Lamb waves traveling in a thin plate. The shape gradient developed in Section 3.4 for the Helmholtz equation may be applied to reconstruct the surface of an unknown hole in the plate, but there are strict limitations on the range of ka values for which this simplified model holds. It is valid for both very large and very small values of ka , and other reconstruction techniques must be used for values between these extremes. Alternatively, the shape gradient could be re-derived by replacing either the governing equations or the boundary conditions with far more complicated expressions. Numerical experiments were used to estimate the ka values for which the solution is valid and to demonstrate the reconstruction.

A laboratory experiment was conducted to provide a data set for reconstruction. The experiment was conducted in Georgia Tech's Quantitative Ultrasonic Evaluation, Sensing and Testing (QUEST) Laboratory. A laser vibrometer was used to collect a set of measurements densely sampled in the vicinity of the scattering hole. This abundance of measurements was intended to support the study of the effects of source/receiver diversity (and the lack thereof). The laser vibrometer is capable of measuring very small displacements normal to the surface of the object under test. The in-plane shear wave components could therefore not be directly observed, although such waves are known to be present even when the incident wave has no shear component. The measured wavefield agrees well with the theory describing the propagation and scattering of elastic waves in thin plates. While the experimental data are of very high quality, the ka values necessary for the shape reconstruction did not match those for which significant energy was transmitted into the plate.

5.1 Elastic Wave Propagation in Plates

The elastodynamic equations of motion in an isotropic medium, in Cartesian coordinates, are

$$\begin{aligned}\mu\nabla^2 u_x + (\lambda + \mu) \left(\frac{\partial^2 u_x}{\partial x^2} + \frac{\partial^2 u_y}{\partial x \partial y} + \frac{\partial^2 u_z}{\partial x \partial z} \right) &= \rho \frac{\partial^2 u_x}{\partial t^2} \\ \mu\nabla^2 u_y + (\lambda + \mu) \left(\frac{\partial^2 u_x}{\partial y \partial x} + \frac{\partial^2 u_y}{\partial y^2} + \frac{\partial^2 u_z}{\partial y \partial z} \right) &= \rho \frac{\partial^2 u_y}{\partial t^2} \\ \mu\nabla^2 u_z + (\lambda + \mu) \left(\frac{\partial^2 u_x}{\partial z \partial x} + \frac{\partial^2 u_y}{\partial z \partial y} + \frac{\partial^2 u_z}{\partial z^2} \right) &= \rho \frac{\partial^2 u_z}{\partial t^2}\end{aligned}\quad (5.1)$$

where $[u_x \ u_y \ u_z]^T = u$ is the displacement vector, $\lambda = 2\mu\nu/(1 - 2\nu)$ is the Lamé constant, ν is Poisson's ratio, and μ is the shear modulus. This system of three equations can be written more generally and compactly in vector form,

$$\mu\nabla^2 u + (\lambda + \mu)\nabla\nabla \cdot u = \rho\ddot{u}, \quad (5.2)$$

or using summation notation

$$\mu u_{i,jj} + (\lambda + \mu)u_{j,ji} = \rho\ddot{u}_i. \quad (5.3)$$

Compared to propagation in fluids, elastic wave propagation is exceptionally complicated, even in cases where the solid medium is isotropic and the geometry apparently simple.

When the medium is infinite in extent, the traveling waves are known as bulk, or body, waves. Body waves exist as either compressional (longitudinal) or shear (transverse) waves, and the respective propagation speeds c_L and c_T are independent of frequency, being purely functions of the material properties of the medium,

$$c_L^2 = \frac{\lambda + 2\mu}{\rho} \quad (5.4a)$$

$$c_T^2 = \frac{\mu}{\rho}, \quad (5.4b)$$

where ρ is the density of the medium.

The presence of boundaries greatly influences the nature of the propagation, and under these conditions the term ‘guided waves’ is used. Solids having one free surface (that is, infinite half-spaces) support surface, or Rayleigh, waves whose amplitude decays exponentially with depth. These waves are of broad interest because they spread cylindrically, not spherically, and therefore carry energy over long distances. Like body waves, Rayleigh waves are nondispersive and their propagation speed [86] is expressed approximately as

$$c_R \approx c_T \frac{0.87 + 1.12\nu}{1 + \nu}. \quad (5.5)$$

Surface waves are often the most destructive component of earthquakes, and they can travel around the earth multiple times before extinguishing.

The presence of both upper and lower boundaries, resulting in an infinite plate geometry, gives rise to Lamb waves. Lamb wave propagation is far more complicated than either body or Rayleigh waves. The plate behaves as a waveguide supporting an infinite number of modes. With the exception of the zero-order modes each has a particular cutoff frequency. The compressional and shear waves interact in a manner that produces two kinds of modes, symmetric (S) and antisymmetric (A). As implied by their names, the particle displacement field for S modes is symmetric about the midplane of the plate, while that of the A modes is antisymmetric. The S_0 mode is a special case, being similar to a purely extensional wave when the wavelength is much larger than the plate thickness.

The propagation speeds for these dispersive S and A Lamb wave modes are described by the Rayleigh-Lamb frequency equations:

$$\frac{\tan(ph)}{\tan qh} = -\frac{4k^2 pq}{(q^2 - k^2)^2} \quad \text{for symmetric modes} \quad (5.6a)$$

$$\frac{\tan(qh)}{\tan ph} = -\frac{(q^2 - k^2)^2}{4k^2 pq} \quad \text{for antisymmetric modes,} \quad (5.6b)$$

where

$$p^2 = \frac{\omega^2}{c_L^2} - k^2 \quad \text{and} \quad q^2 = \frac{\omega^2}{c_T^2} - k^2$$

and h is the half-thickness of the plate. These equations were used to design a guided wave experiment such that only the first of each kind of mode, A_0 and S_0 , are supported over the range of frequencies excited by a set of ultrasonic transducers attached to the plate.

Since the measured quantity in elastic wave experiments is often the normal displacement u_z , Equations 5.1 can be manipulated to eliminate u_x and u_y . The result is a sixth-order partial differential equation, which would be extremely difficult to use for deriving the shape gradient. Fortunately, a convenient alternative exists. The displacements can be written in terms of derivatives of potential functions, which in turn satisfy uncoupled wave equations [87]. Suppose that the displacement vector can be written as the following decomposition

$$u = \nabla\phi + \nabla \times \psi \tag{5.7}$$

where ϕ and ψ are scalar and vector potentials. The vector ψ has two components, the horizontal and vertical shear potentials ψ_h and ψ_v . Substituting this equation into 5.1 and noting that $\nabla \cdot \nabla\phi = \nabla^2\phi$ and $\nabla \cdot \nabla \times \psi = 0$ gives

$$\nabla \left[(\lambda + 2\mu)\nabla^2\phi - \rho\ddot{\phi} \right] + \nabla \times \left[\mu\nabla^2\psi - \rho\ddot{\psi} \right] = 0. \tag{5.8}$$

This equation is satisfied if the following hold

$$\begin{aligned} \nabla^2\phi &= \frac{1}{c_L^2}\ddot{\phi} \\ \nabla^2\psi &= \frac{1}{c_T^2}\ddot{\psi}. \end{aligned} \tag{5.9}$$

These uncoupled potential wave equations are much simpler to work with, and mean that the shape gradient developed in Chapter 3 may carry over to the elastodynamic problem. However, it is important to remember that coupling and high-order

derivatives remain embedded in the boundary conditions. Thus the ability to employ the shape gradient of Section 3.4 depends on finding conditions under which simple boundary conditions apply to the potential itself. While this is not possible in general, the high- and low-frequency regimes permit suitable approximations to be made. This is discussed further below.

5.2 Elastic Wave Scattering in Plates

The experiment conducted for this research is specifically concerned with waves in plates. There exist a number of plate-specific theories that vary in their complexity and regimes of validity. Plate theories employ a variety of approximations to reduce the equations of motion from three to two dimensions. The use of these is largely motivated by problems in which the plate is finite in extent [87], although in the present case the experiment was contrived such that the influence of the plate boundaries was minimized.

A useful model for wave propagation and scattering from circular inclusions can be found in Chapter 3 of Mow and Pao [48]. That model is developed for plane strain conditions in an infinite medium, but it can be converted to plane stress by an appropriate transformation of the constant coefficients in the governing equations. Plane stress conditions require the principal stress σ_{zz} normal to the plate to be zero. For elastodynamic waves, a plate satisfies conditions of plane stress (and is said to be thin) when its thickness is small relative to the propagating wavelength. The plane strain solution is exact for an infinite medium. However, the conversion to plane stress and application to thin plates is an approximation, akin to Poisson plate theory [87], where the waves being modeled are purely extensional (or longitudinal). The experiment was designed to satisfy thin-plate conditions. The model based on Mow and Pao used for this research cannot account for dispersion since it is a rescaling of

a model describing bulk wave in an infinite medium, which is necessarily nondispersive (see Equations (5.4)). However, appropriate wave speeds can be found from the equations described in the previous section. A similar reasoning was used by Grahn to model S_0 and A_0 scattering [88].

The incident wave is assumed to be a compressional wave. For this experiment, the S_0 mode approximates a purely compressional wave when the wavelength is larger than the plate's thickness. A detailed analysis of the displacement profile across the plate thickness can be found in Chapter 8 of Rose [47]. The boundary conditions at the surface of the inclusion permit coupling such that the scattered field comprises both compressional and horizontal shear waves. There is no vertical shear component in the plate when the incident wave is compressional.

The compressional and horizontal shear displacement potentials, ϕ and ψ , are used in a Fourier expansion to find the total field. The origin for the cylindrical coordinates (r, θ) is placed at the center of the scattering hole. The incident compressional wave is often modeled as being planar, but a cylindrical wave source is a better representation of the experimental conditions, where the source was in relatively close proximity to the scatterer. The incident potential can be expressed using the following expansion in cylindrical harmonics:

$$\phi^i = \phi_0 \sum_{n=0}^{\infty} \Lambda_n J_n^{(1)}(\alpha r) \cos(n\theta) e^{-i\omega t}, \quad (5.10)$$

where

$$\epsilon_n = \begin{cases} 1 & \text{if } n = 0 \\ 2 & \text{if } n \geq 1, \end{cases} \quad (5.11)$$

and α is the compressional wavenumber. The coefficient Λ_n equals $\epsilon_n i^n$ if the incident wave is planar, and $\Lambda_n = \epsilon_n (-1)^n H_n^{(1)}(\alpha r_0)$ if the incident wave is cylindrical. In the latter case, r_0 is the distance from the center of the scattering hole to the source.

The scattered compressional and horizontal shear wave potentials are:

$$\phi^s = \sum_{n=0}^{\infty} A_n H_n^{(1)}(\alpha r) \cos(n\theta) e^{-i\omega t} \quad (5.12)$$

$$\psi_h^s = \sum_{n=0}^{\infty} B_n H_n^{(1)}(\beta r) \sin(n\theta) e^{-i\omega t}, \quad (5.13)$$

where A_n and B_n are the unknown Fourier expansion coefficients. The scattered compressional waves are even in θ , while the scattered shear waves are odd. The two unknown Fourier coefficients are found by imposing a pair of boundary conditions at the surface of the hole. Since the hole represents a free surface, both σ_{rr} and $\sigma_{r\theta}$ must be zero. These stress components are expressed in terms of the potentials as

$$\sigma_{rr} = \lambda \nabla^2 \phi + 2\mu \left[\frac{\partial^2 \phi}{\partial r^2} + \frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial \phi}{\partial \theta} \right) \right] \quad (5.14)$$

$$\sigma_{r\theta} = \mu \left\{ 2 \left[\frac{1}{r} \frac{\partial^2 \phi}{\partial \theta \partial r} - \frac{1}{r^2} \frac{\partial \phi}{\partial \theta} \right] + \left[\frac{1}{r^2} \frac{\partial^2 \psi_h}{\partial \theta^2} - r \frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial \psi_h}{\partial r} \right) \right] \right\}. \quad (5.15)$$

Substituting the frequency-domain version of (5.9) gives the following system of equations for the Fourier coefficients

$$\begin{bmatrix} \varepsilon_{11}^{(3)} & \varepsilon_{12}^{(3)} \\ \varepsilon_{41}^{(3)} & \varepsilon_{42}^{(3)} \end{bmatrix} \begin{bmatrix} A_n \\ B_n \end{bmatrix} = \begin{bmatrix} -\phi_0 \Lambda_n \varepsilon_{11}^{(1)} \\ -\phi_0 \Lambda_n \varepsilon_{41}^{(1)} \end{bmatrix} \Big|_{r=a} \quad (5.16)$$

The coefficient Λ_n is defined above, and the terms $\varepsilon_{mn}^{(p)}$ are given in the Chapter 3 Appendix of [48]. The latter represent the contributions of the potentials to a given stress component. The principal stresses can be computed everywhere on the plate according to:

$$\sigma_{rr} = \frac{2\mu}{r^2} \sum_{n=0}^{\infty} \left[\phi_0 \Lambda_n \varepsilon_{11}^{(1)} + A_n \varepsilon_{11}^{(3)} + B_n \varepsilon_{12}^{(3)} \right] \cos(n\theta) e^{-i\omega t} \quad (5.17)$$

$$\sigma_{\theta\theta} = \frac{2\mu}{r^2} \sum_{n=0}^{\infty} \left[\phi_0 \Lambda_n \varepsilon_{21}^{(1)} + A_n \varepsilon_{21}^{(3)} + B_n \varepsilon_{22}^{(3)} \right] \cos(n\theta) e^{-i\omega t}. \quad (5.18)$$

Hooke's law is then used to find the principal strain ε_{zz} normal to the plate's midplane

$$\varepsilon_{zz} = \frac{1}{E} [\sigma_{zz} - \nu(\sigma_{rr} + \sigma_{\theta\theta})], \quad (5.19)$$

where $E = \mu(3\lambda + 2\mu)/(\lambda + \mu)$ is the elastic (or Young's) modulus. From this, the surface displacement that would be measured is given by $h\varepsilon_{zz}$ where h is the half-thickness of the plate.

The plate is assumed to be in a state of plane stress, so only σ_{rr} and $\sigma_{\theta\theta}$ are nonzero in (5.19). This fact permits the circumvention of an otherwise significant complication. While the displacement potentials allow for easy modeling of wave propagation, they are generally not linked to measurable quantities via simple expressions. The shape gradient framework of Section 3.4 cannot be used when higher-order derivatives are needed to relate the potential to measurements. Fortunately, it can be shown that the potential ϕ is proportional to the vertical displacement in the present context. The principal stress σ_{rr} appears in (5.14), and the stress $\sigma_{\theta\theta}$ is given by:

$$\sigma_{\theta\theta} = \lambda\nabla^2\phi + 2\mu \left[\frac{1}{r} \left(\frac{\partial\phi}{\partial r} + \frac{1}{r} \frac{\partial^2\phi}{\partial\theta^2} \right) + \frac{1}{r} \left(\frac{1}{r} \frac{\partial\psi_h}{\partial\theta} - \frac{\partial^2\psi_h}{\partial r\partial\theta} \right) \right]. \quad (5.20)$$

Adding these together and using the definition of the Laplacian operator ∇^2 in cylindrical coordinates results in

$$(\sigma_{rr} + \sigma_{\theta\theta}) = -2k^2(\lambda + \mu)\phi. \quad (5.21)$$

Combining this result with Hooke's law, the measured displacement normal to the plate's midplane is

$$u_z = \frac{2h\nu k^2}{E} (\lambda + \mu) \phi, \quad (5.22)$$

therefore establishing proportionality between the measured quantity u_z and the potential function ϕ .

There is one additional consideration that must be addressed to determine if the shape gradient based on the Helmholtz equation can be applied to the potential function. Table 3.1 shows the boundary conditions that may be imposed on the forward model to obtain the shape gradient. If one of these does not hold, then the derivation likely cannot be carried out without reformulating the problem completely.

It was stated previously that the conditions on the boundary of the scattering hole are coupled and expressed in terms of stresses, not the potential function. Therefore, the shape gradient of Section 3.4 can only be applied if there are circumstances under which it can be said that the boundary conditions of Table 3.1 apply to the potential function itself. This is indeed the case if the incident wavelength is either very long or very short ($ka \ll 1$ or $ka \gg 1$, respectively).

Recall that the normal stress σ_{rr} is equal to zero at the face of the hole. From (5.21) it can be seen that the potential ϕ is proportional to the principal stress $\sigma_{\theta\theta}$ on this boundary

$$\phi|_{r=a} = -\frac{\sigma_{\theta\theta}}{2k^2(\lambda + \mu)}. \quad (5.23)$$

Pao [89] describes the long- and short-wavelength limits for scattering from a circular hole in terms of $\sigma_{\theta\theta}$. For very low ka the conditions at the hole approach that of static loading:

$$\sigma_{\theta\theta}|_{r=a} \propto (k^2 - 1 - 2\cos(2\theta)), \quad (5.24)$$

while for very high ka , the short waves effectively encounter a stress-free surface such that

$$\sigma_{\theta\theta} \approx 0. \quad (5.25)$$

From (3.23) the condition needed is that the derivative of ϕ with respect to the gradient descent time is zero, or $\phi_t = 0$. This is satisfied by both of the limiting cases. For very small ka , the value of ϕ at $r = a$ isn't a function of the hole radius or position, so it doesn't change during the gradient descent. For very large ka , ϕ simply approaches zero.

5.3 Experimental Setup and Signal Processing

The experimental measurements were made using a square aluminum 6061-T6 plate [90] that was 1,219 mm (48 in.) on a side and whose thickness was 2.286 mm (0.090 in.). Six piezoelectric ultrasonic transducers were bonded to the plate in an approximate

Table 5.1: Positions of the six piezoelectric transducers bonded to the experimental plate. The origin is located at the center of the plate.

Sensor No.	x (mm)	y (mm)
1	232	74
2	53	238
3	-177	168
4	-228	-68
5	-62	-236
6	167	-164

ring configuration. These were 7 mm diameter Steiner & Martins model SMD07T05R411, whose resonant frequency is 300 kHz. The positions are given in Table 5.1. An exact ring distribution was avoided to break up the symmetry of the source/measurement locations. The radius of the transducer ring was chosen such that scattering from a hole near the center of the plate could be recorded by any transmit/receive pair before reflections arrived from the plate edges. This placement of the sensors allows the edge reflections to be time gated and the plate treated as being infinite in extent. However, for the experiment described below the piezoelectric transducers were used only as sources, and a Polytec OFV-525/-5000-S scanning laser vibrometer provided a set of measurements on a dense grid near the center of the plate.

A 5 mm hole was drilled above and to the right of the plate’s center by 20 mm in each direction. This hole serves as the unknown target whose position and radius is to be reconstructed. The experimental plate configuration is shown in Figure 5.1. The serrated regions at the edges indicate the placement of damping material used to suppress edge reflections and shorten the time required for the plate to ring down between measurements. This material was Skandia AeroDamp SK-8240PSA damping sheet whose primary commercial application is noise suppression in aircraft cabins. While

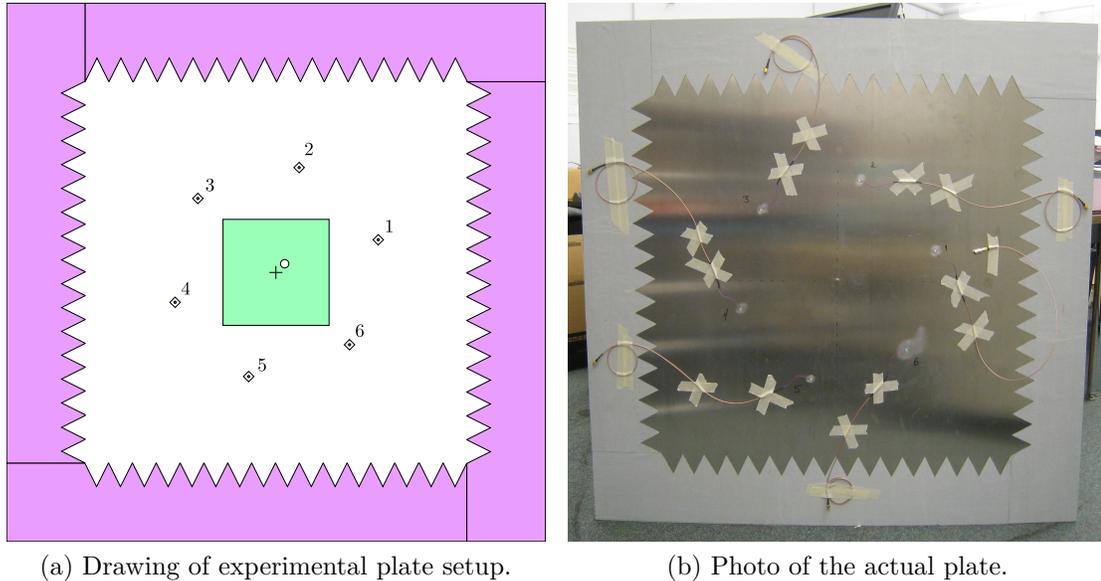


Figure 5.1: The drawing at left shows the components used in the plate experiment. The six sources are numbered 1-6, and their positions are indicated by the diamonds. The photo at right shows the plate as it was constructed. Also visible in the photo are the cables attached to the source transducers.

optimized for audible frequencies, this material worked well for its intended purpose in the experiment. The square in the center of the drawing is the 240 mm \times 240 mm region of support for the laser vibrometer measurements. Measurements were made in 2 mm steps, yielding a 121 \times 121 grid of spatial samples. The small circle above and to the right of the plate center is the 5 mm circular hole used as the target. The drawing is to scale except for this hole, which is shown four times larger so that it can be easily seen.

The transducers were excited independently using a linearly-swept frequency-modulated (LFM), or chirp, signal. The LFM is the most commonly used waveform in radar and sonar applications because it can provide both fine range resolution and good SNR [91, 92] without having to transmit large amounts of instantaneous power. The advantage of this waveform for nondestructive testing purposes is that any of the frequencies can be isolated in postprocessing, thus obviating the need for making separate measurements for each desired frequency [45]. The chirp used swept

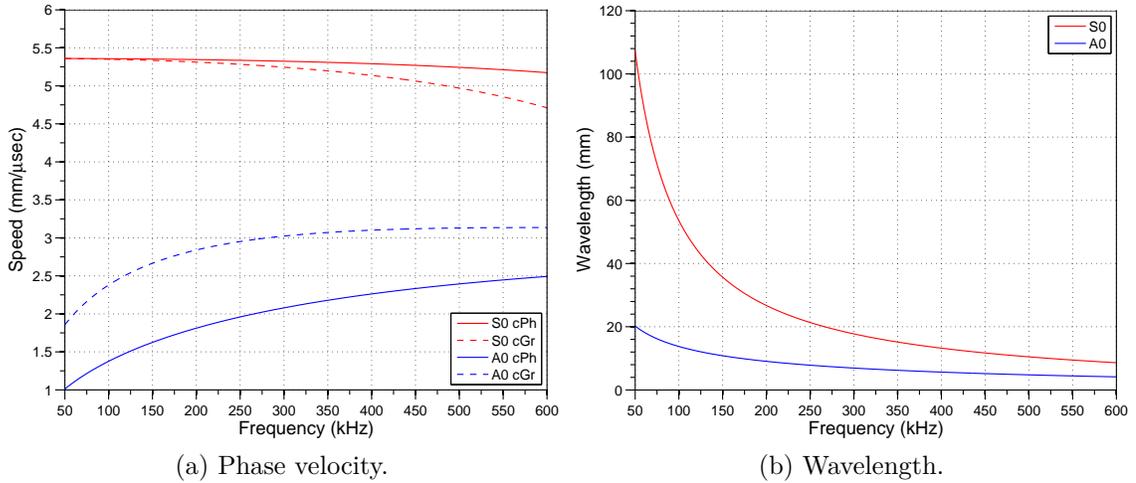


Figure 5.2: Phase and group velocities and the corresponding wavelengths for S_0 (red) and A_0 (blue) Lamb wave modes.

from 50–600 kHz over a pulse length of 180 μs . The pulse repetition interval (PRI) was 2 ms. This PRI, combined with the damping material, provided sufficient time between transmissions for the waves in the plate to diminish to negligible levels.

The combination of frequencies and hole radius was chosen to provide ka values both less than and greater than one. Furthermore, the frequencies and plate thickness were chosen so as to permit only the propagation of the S_0 and A_0 Lamb wave modes. These are the lowest-order symmetric and antisymmetric modes, respectively [93, 47], and they correspond approximately to ideal compressional and flexural waves. The phase and group velocities and wavelengths are shown in Figure 5.2 for the plate thickness and range of frequencies used in this experiment. Because the propagation speeds for S_0 and A_0 waves are substantially different, the measured wavefield can be filtered spatially to retain only the mode of interest. In this case the S_0 mode was chosen because it travels faster, thus arriving in the measured scene before any other waves. Retaining only the S_0 wave also ensures that the experimental incident wave closely matches the idealized theory. For reference, values of the wavelength at several frequencies spanning the transmitted band are collected in Table 5.2.

The scanner stepped a raster over the 240 mm square box shown in Figure 5.1

Table 5.2: Phase speed and wavelength for the S_0 and A_0 modes at several transmitted frequencies. Also shown is the value of ka for $a = 2.5$ mm.

f (kHz)	c_{S_0} (m/s)	λ_{S_0} (mm)	ka_{S_0}	c_{A_0} (m/s)	λ_{A_0} (mm)	ka_{A_0}
50	5360	107	0.15	1012	20	0.78
100	5357	54	0.29	1376	14	1.1
150	5352	36	0.44	1625	11	1.5
200	5345	27	0.59	1812	9	1.7
250	5336	21	0.74	1959	8	2.0
300	5324	18	0.89	2079	7	2.3
350	5309	15	1.0	2179	6	2.5
400	5291	13	1.2	2262	6	2.8
450	5269	12	1.3	2333	5	3.0
500	5243	10	1.5	2394	5	3.3
550	5211	9	1.7	2446	4	3.5
600	5173	9	1.8	2492	4	3.8

in 2 mm increments. At each position (x_m, y_n) , a single source was excited and the plate displacement was recorded until the waves were effectively extinguished. This procedure was repeated for sources 1, 3, and 5. The resulting collection of all time series $S_{mn}(x_m, y_n, t)$ is known in the radar literature as a data cube. The data cubes can be filtered in any dimension to obtain a desired set of information. Since the propagation is linear, the data cubes may be added together to study the effect of using multiple sources simultaneously. The signal processing steps applied to the collected data are:

1. Deconvolve the transmitted FM chirp (also known as matched filtering).
2. Apply a narrowband temporal filter to isolate a single frequency from the transmitted band.

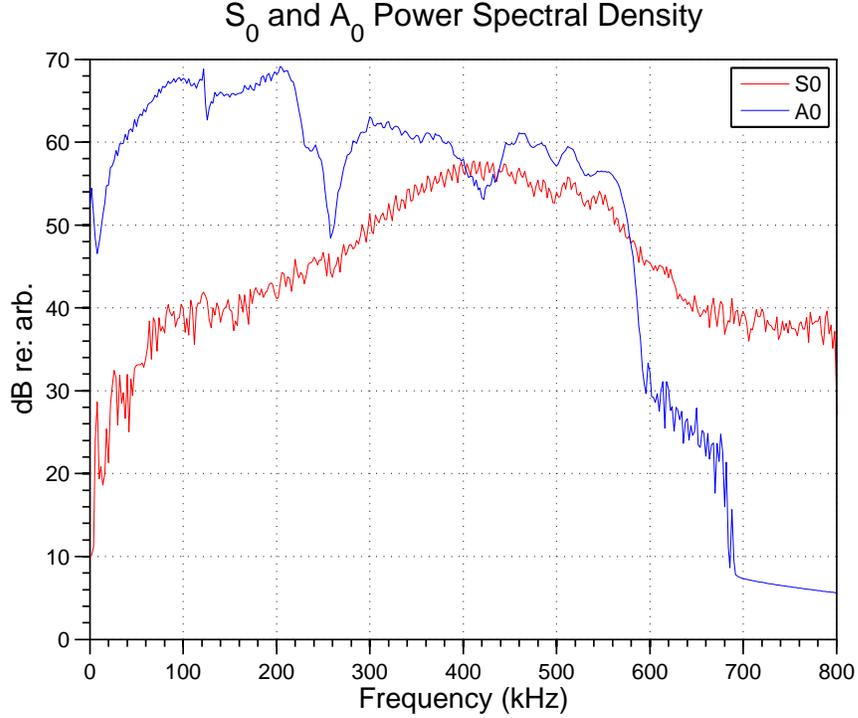


Figure 5.3: Relative power spectral density for S_0 and A_0 modes.

3. Although the data now have a single temporal frequency, ω , there may be multiple spatial frequencies $k = \omega/c$ because the modes travel with different speeds, as shown in Figure 5.2. A two-dimensional spatial bandpass filter is applied to reject all but the desired mode [94].
4. The spatial envelope of the wavefield is obtained from the filtered data cube by taking the maximum value of the magnitude of the time series at each measurement location.

Figures 5.3 and 5.4 give an idea of the relative strengths of the S_0 and A_0 modes over the transmitted bandwidth. Figure 5.3 shows the temporal spectrum of both modes. The A_0 mode is strong over a wide span of frequencies, with notable dropouts around 250 kHz and 435 kHz. Meanwhile, the S_0 mode only has appreciable energy between approximately 250 kHz and 625 kHz. This discrepancy between S_0 and A_0 excitation using bonded piezoelectric wafer transducers is known in the literature as

mode selectivity. For example Nieuwenhuis et al. discuss this effect and compare finite-element predictions with laboratory measurements [95].

The relative strengths of these modes is visualized in Figure 5.4. The location of Source 1 is $(x, y) = (232, 74)$ mm, and the time series signal at each location was filtered to exclude all but a 5-cycle tone burst at the center frequency indicated for each subplot. The S_0 mode is relatively weak for frequencies up to 200 kHz. The S_0 mode is stronger above this frequency, and scattering from the hole located at $(x, y) = (20, 20)$ mm becomes visible. Meanwhile, the A_0 mode is strong at all of the frequencies shown, in keeping with Figure 5.3.

5.4 Model-Data Comparison

Several aspects of the wavefield observed in the plate were compared to the theoretical predictions. The first of these is the dispersion of the S_0 and A_0 modes. This comparison was made by extracting a slice of the data cube corresponding to the radial line starting at the location of Source 1 and passing through the center of the scan area. The data were obtained by interpolating the wavefield at any given time onto the red line shown in Figure 5.5. The resulting signal is denoted as $S(r, t)$, where r is radial distance along the slice. These measurements were made before the hole was drilled into the plate, so no scattering is observed.

Taking the two-dimensional Fourier transform of $S(t, r)$ gives $S(f, k)$, providing access to a domain in which the S_0 and A_0 modes can be clearly distinguished. The modes were separated by masking out regions of the (f, k) support that do not correspond to the desired mode. These masked spectra are shown in Figure 5.6. The theoretical dispersion curves are plotted for each mode, and were computed using Equations (5.6). The measured and predicted dispersion in the spectral domain are in near perfect agreement.

The filtered spectra were transformed back to the (t, r) domain and are plotted in

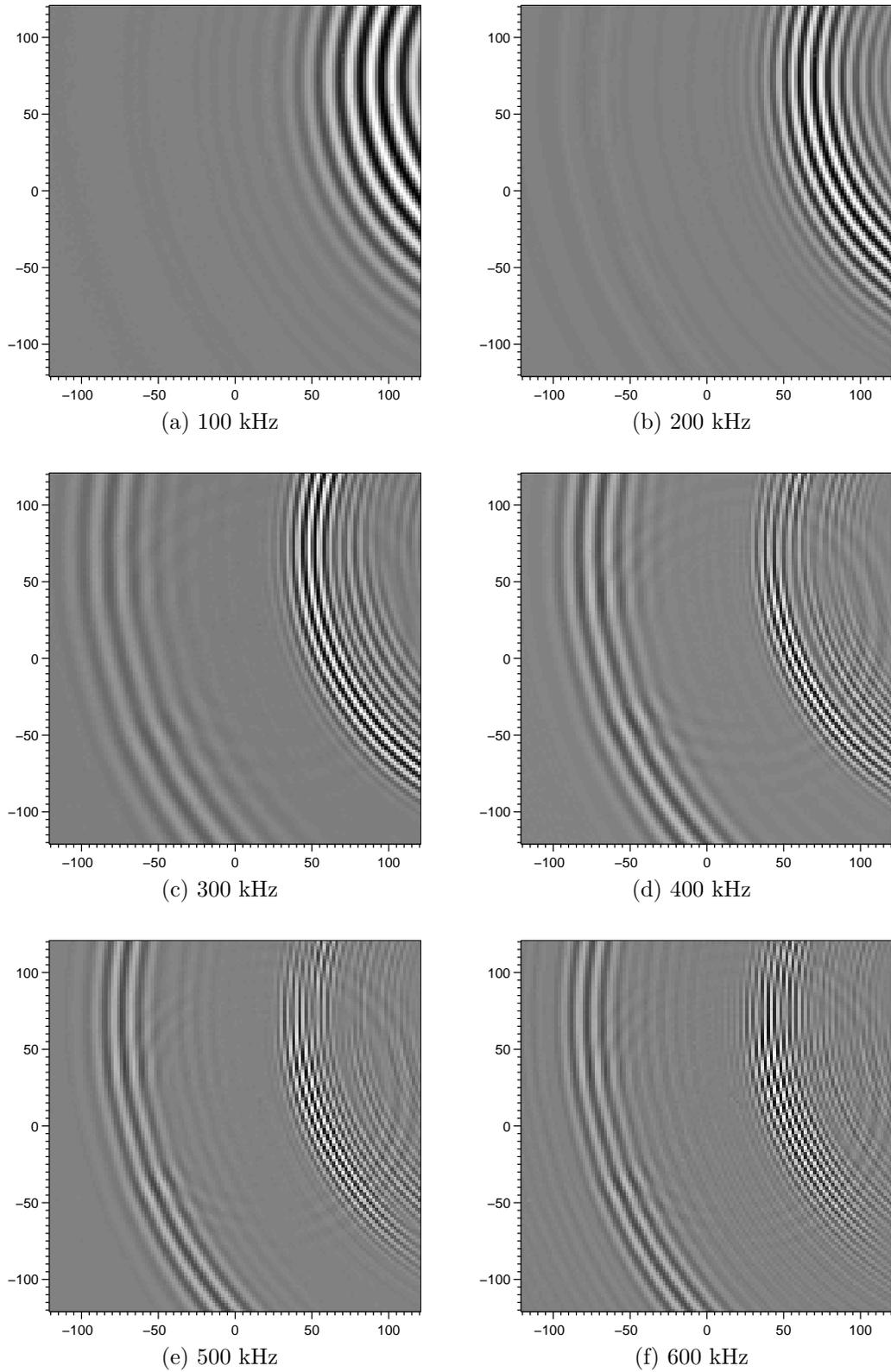


Figure 5.4: Measured wavefields at six different frequencies within the transmitted bandwidth. The time index of each image was chosen such that the faster traveling S_0 wave packet had just passed the center of the measurement area. The axes are labeled in units of mm.

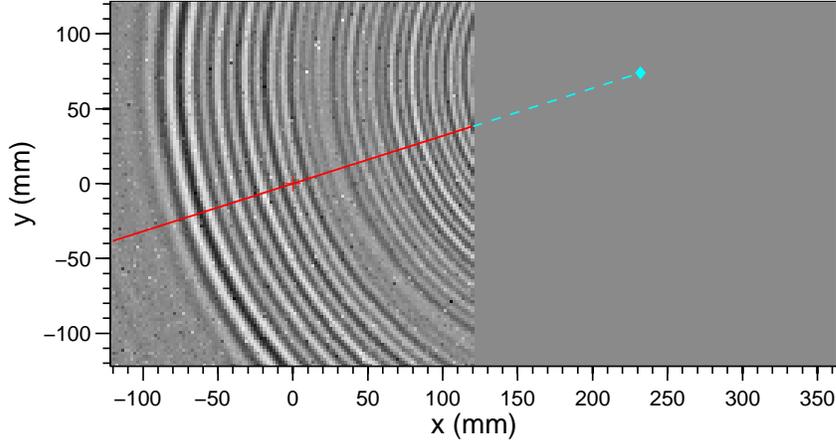


Figure 5.5: A slice through the wavefield was extracted along a radial line originating at Source 1 and passing through the center of the scan area. The red segment shows the domain of the slice, and the cyan extension reaches to the location of Source 1. The wavefield shown in the figure corresponds to $t = 150 \mu\text{sec}$.

Figure 5.7. The theoretical phase and group velocities were used to overlay characteristic curves that can be compared to the observed wavefronts. The characteristics are plotted for frequencies ranging from 50 kHz to 600 kHz in 50 kHz steps (12 total). The characteristics are delayed according to the distance to the source (outside the scan area) and according to the time at which each frequency occurred in the transmitted chirp signal. For example, the 50 kHz characteristic begins at $t = 0$ s, and the 600 kHz line does not leave the source until $t = 180 \mu\text{s}$. The S_0 characteristics are parallel to the paths of the traveling wavefront providing further evidence of the good agreement between theory and measurement. This comparison is made easier by the fact that the S_0 phase and group velocities are very similar. Interpretation of the A_0 data is little more complicated because the phase and group velocities differ significantly, and the amount of this difference varies with frequency. The A_0 portion of the transmitted chirp has a well-developed leading edge wave packet traveling with the 50 kHz group velocity. Close examination of the plot shows that the phase velocity of the wavefronts within the packet travel at the 50 kHz phase velocity (the red characteristic with the shallowest slope).

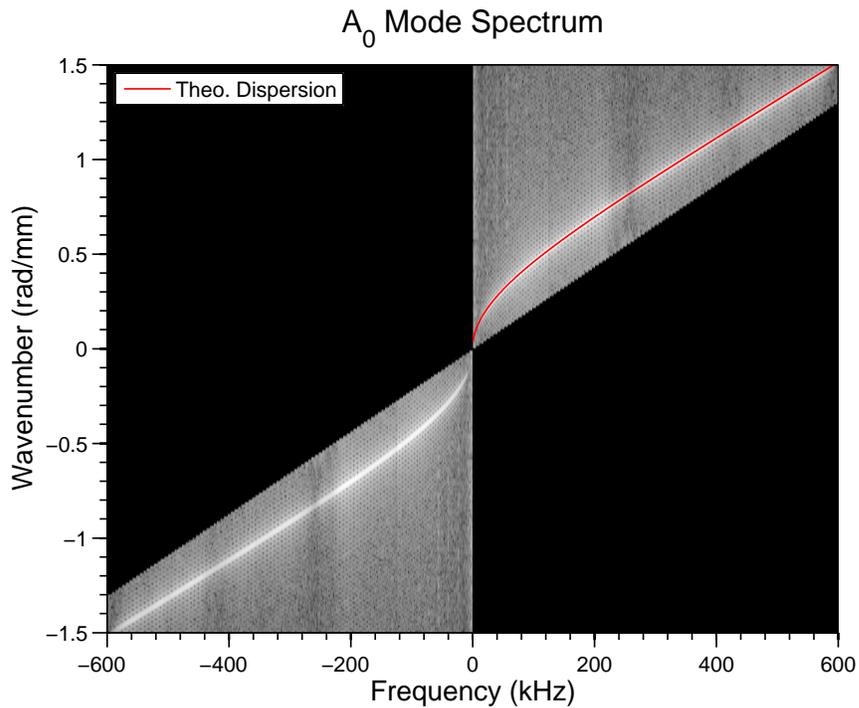
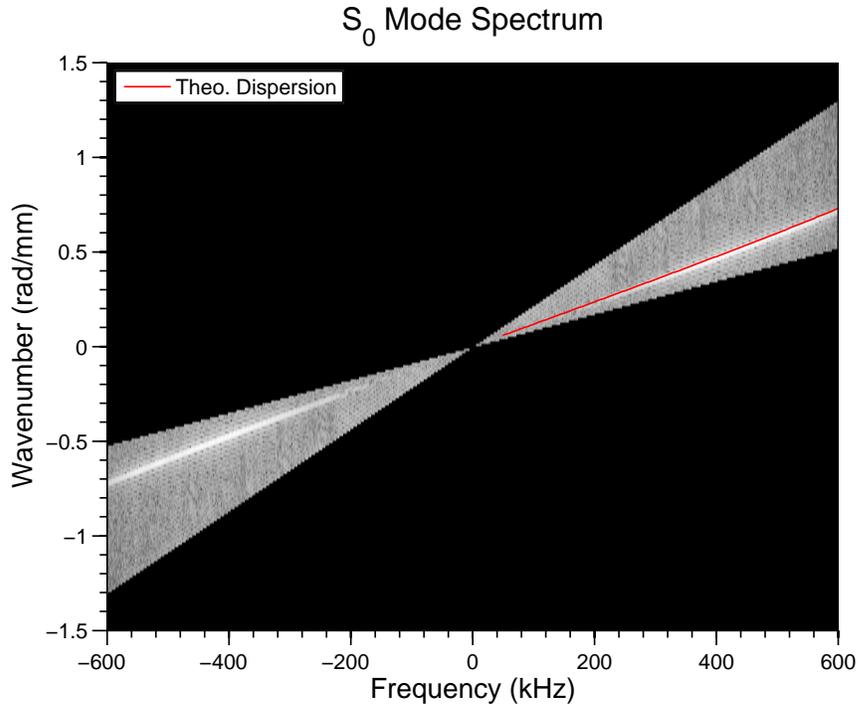
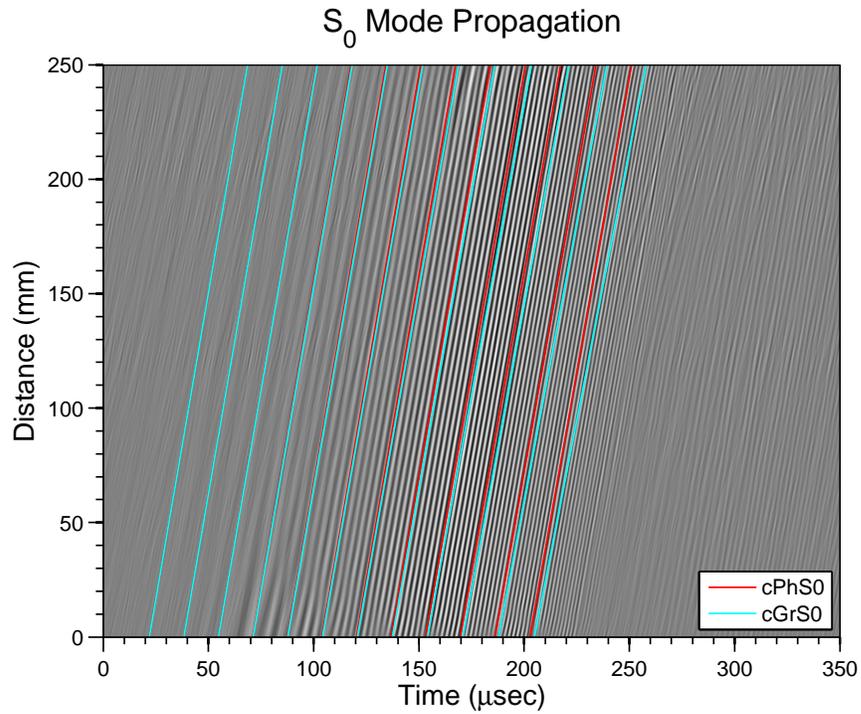
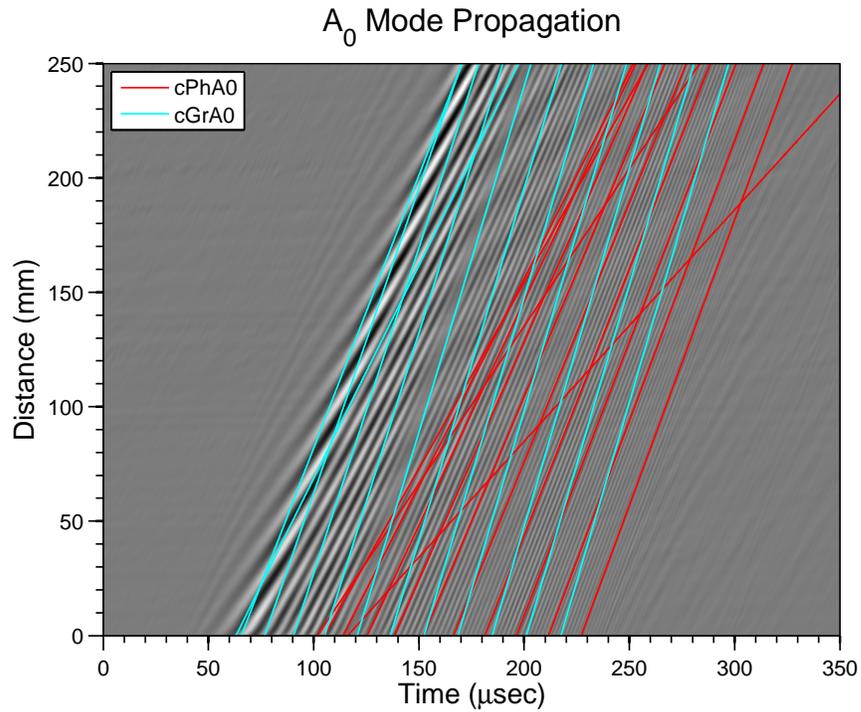


Figure 5.6: The two-dimensional spectrum of the data extracted from the slice shown in Figure 5.5 was filtered to extract only the desired mode of propagation. The red curves indicate the theoretical dispersion curves computed from the Rayleigh-Lamb equations.



(a) Measured S_0 radial propagation.



(b) Measured A_0 radial propagation.

Figure 5.7: S_0 and A_0 mode propagation along the radial line shown in Figure 5.5. The red and cyan lines indicate the characteristic curves associated with the phase and group velocities, respectively.

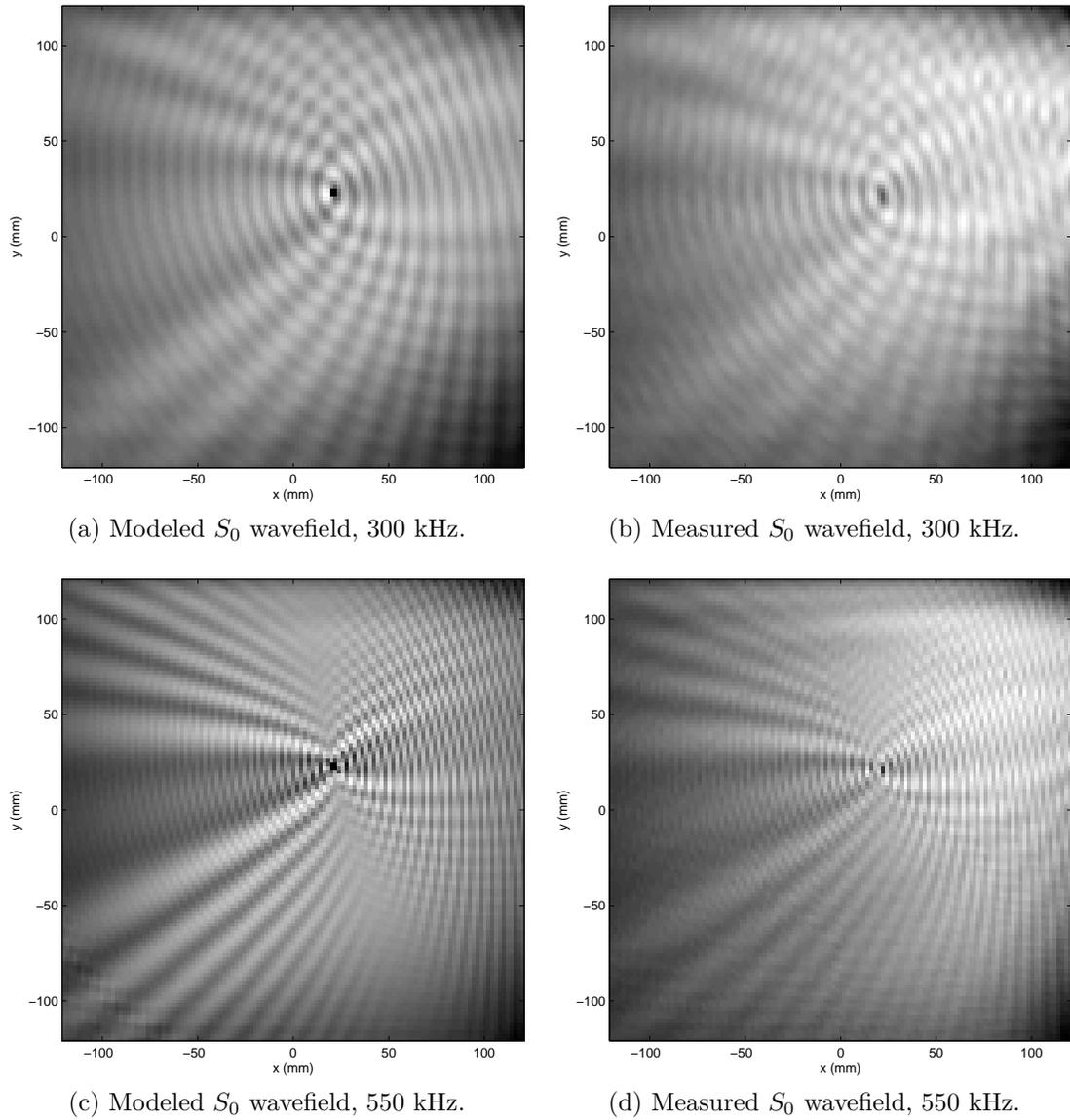


Figure 5.8: Comparison between the wavefield predicted using the Fourier expansion (left) and the experimentally measured wavefield (right). Both wavefields have been spatially filtered to retain only the S_0 mode.

Next, the harmonic expansion scattering model described in Section 5.2 was used to predict the total S_0 wavefield on the plate within the region measured by the laser vibrometer. The results are shown in Figures 5.8 and 5.9 for frequencies of 300 kHz and 550 kHz. The phenomenon visible in these figures is the vertical displacement component of the near-field interference pattern due to the scattered waves interacting with the incident wave. As described earlier, the S_0 mode was isolated by spatially

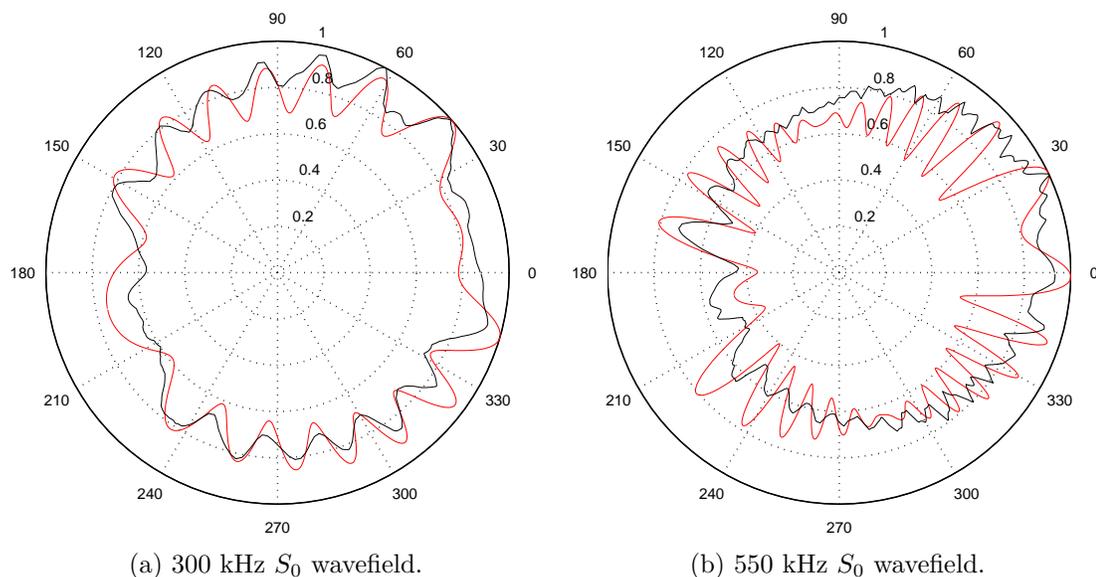


Figure 5.9: Comparison between the wavefield predicted using the Fourier expansion (red) and the experimentally-measured wavefield (black). The curves represent an azimuthal cut through the wavefields shown in Figure 5.8. The cuts are centered on the scattering hole with a radius of 64 mm.

filtering the wavefield for each time sample. Results below 250 kHz were unreliable because of the low amount of S_0 energy present.

Although the wave propagation model is approximate, it exhibits good agreement with the measurements. At higher frequencies, the model should break down as the z -dependence of the stress field becomes significant. At very low frequencies, the model remains valid but becomes difficult to compare to measurements since little S_0 energy was transmitted below 250 kHz. Furthermore, The scattering for low ka values is weak, and thus can be difficult to observe due to measurement noise and any slight levels of reverberation in the plate from earlier source triggers. Regardless of ka , this reverberation may be partially responsible for the lack of null depth in the azimuth cuts shown in Figure 5.9.

5.5 Shape Recovery via Active Surfaces

For suitable values of ka the location and boundary of the hole may be reconstructed using the shape gradient developed in Section 3.4 coupled with the model of Section 5.2. The properties of the reconstruction can be examined by visualizing the behavior of the cost function (or energy landscape) over the region of observation, as was done in Chapter 4. The reconstruction is performed using knowledge of the source(s), measurements made over a region known to surround the target, and a reasonably accurate model for wave propagation and scattering. The algorithm is iterative: At each step the mismatch between the measured and predicted scattering is used to refine an estimate of the target properties. Further details are given in Section 4.2 and in Cook et al. [96].

For the present experiment, any of the points in the scan area can be used for reconstruction. This permits the investigation of the effect of measurement density and geometry. The results presented here employ a ring of measurements similar to that used to derive the plots in Figure 5.9. The spacing of the points around the ring is $\lambda/8$, where λ is the wavelength of the S_0 mode at any given frequency. Sources 1, 3, and 5 were active, and Sources 2, 4, and 6 were not.

The frequency diversity of the experiment allows the performance of the algorithm to be examined as a function of ka . The available bandwidth can also be used to resolve ambiguities that might otherwise trap the reconstruction in a local minimum. This is a well-grounded approach, and is similar to multi-frequency techniques used to resolve ambiguities in certain radar signal processing applications.

It was shown earlier that the shape gradient derived from the potential function has a simple form and is easy to work with, but it also imposes rather strict conditions on the range of ka for which it is suitable. The experiment conducted for this research does not correspond to this range of ka , so the shape gradient does not behave as well as desired. Numerical simulations of the active surface reconstruction were performed

at 10 kHz, 25 kHz, 50 kHz, 75 kHz, and 100 kHz since these are the frequencies at the end of the band closest to ka values for which the shape gradient is expected to be valid. Results are summarized in Figures 5.10–5.14.

The cost surface was computed using only the scattering model from Section 5.2, which was previously shown to agree well with the measured wavefield. At higher frequencies these cost surfaces contain a strong global minimum at the true location of the hole, as well as some weaker local minima. The local minima appear to fall approximately on half-wavelength radial intervals centered on the true location of the hole. If this is the case, then a practical implementation of the reconstruction should employ the dispersion relationship to determine a set of wavelengths (and therefore frequencies) for which the local minima of the cost functions do not coincide. This information could also be used to establish the order of a frequency-hopping approach to finding the global minimum.

The bottom rows of Figures 5.10–5.14 contain vector fields indicating the direction of the gradient of the cost function computed two ways. The left-hand vector field is the result of computing the gradient by finite-differencing the cost function itself. Therefore, this gradient depends only on the forward scattering model, which was shown to agree closely with the measurements. The right-hand gradient field results from the shape gradient given by Equation (3.25), and it depends on the accuracy of the approximate boundary conditions imposed on the displacement potential. Ideally, these two vector fields would be in close agreement, but they diverge for ka exceeding approximately 0.1. This fact, combined with the lack of low-frequency S_0 energy, means that the particular shape gradient derived for the reconstruction is not well-matched to the usable frequency regime of the experiment.

Despite this result, there may be applications for which the existing shape gradient is a good approximation. Unfortunately, deriving the exact shape gradient for this experiment is likely to be very difficult, if not impossible because of the high-order

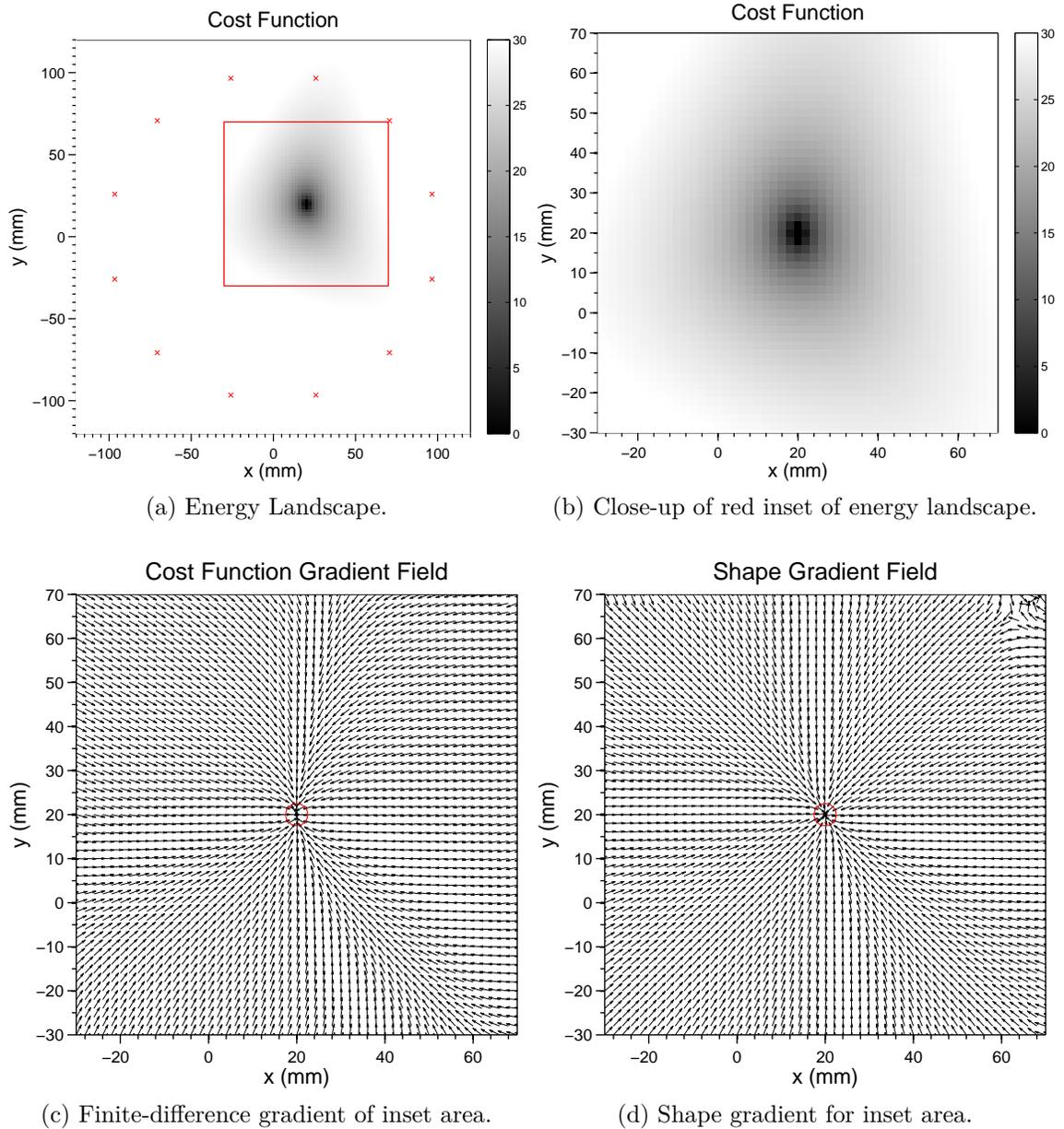


Figure 5.10: A numerical simulation of shape gradient reconstruction was conducted for S_0 waves at 10 kHz with 12 measurement points forming a ring about the origin of the scan area. Sources 1, 3, and 5 were included in the simulation, but are not shown. The top row shows the energy landscape (cost function) and a close-up of the region around the true hole location at $(x, y) = (20, 20)$ mm. The red markers indicate the measurement locations. The bottom row shows the finite difference gradient field of the cost function and that computed from the shape gradient. The vectors all have unit magnitude and are intended to show direction only. The red circle indicates the outline of the 5 mm hole in the plate.

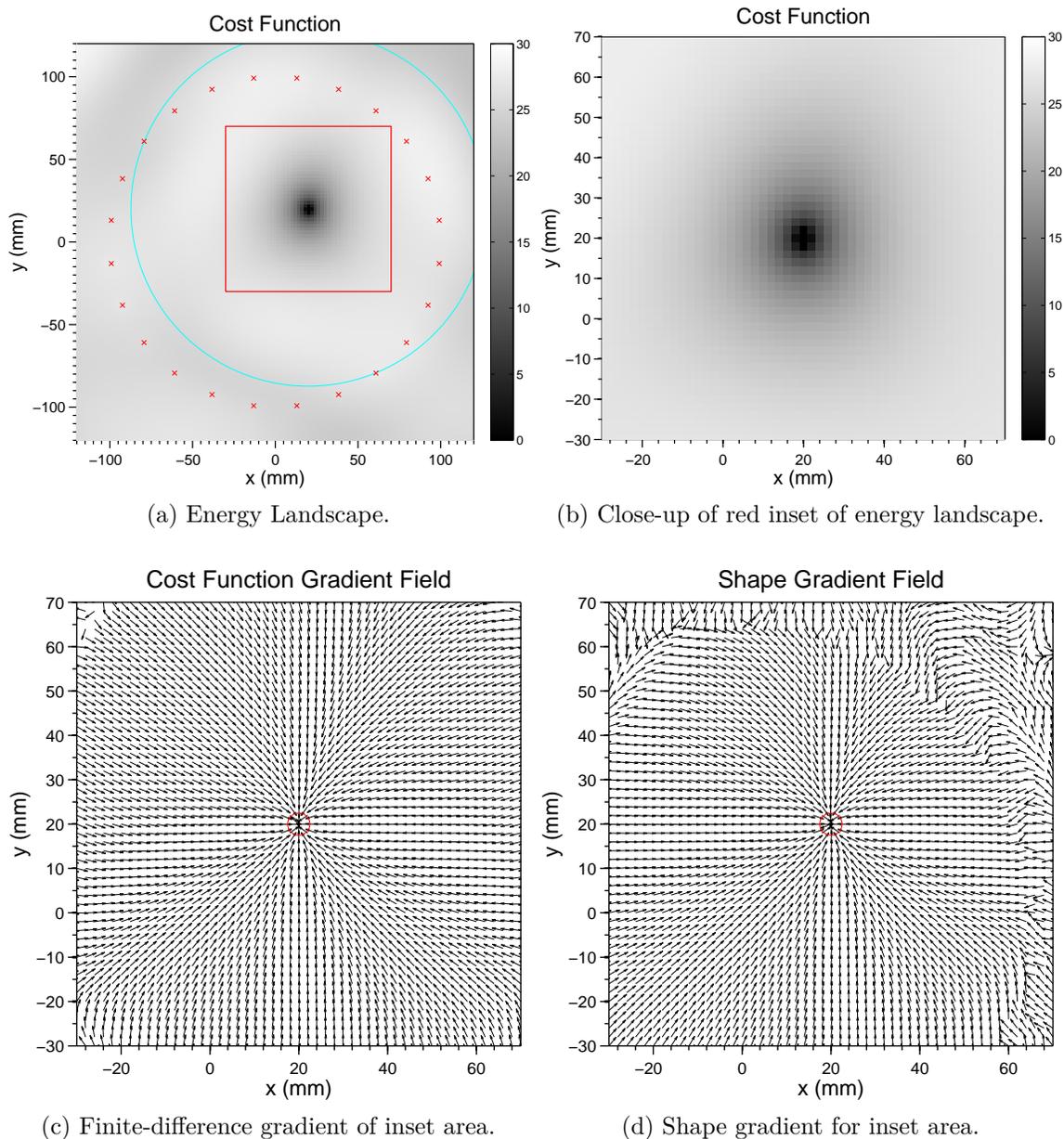


Figure 5.11: A numerical simulation of shape gradient reconstruction was conducted for S_0 waves at 25 kHz with 24 measurement points forming a ring about the origin of the scan area. Sources 1, 3, and 5 were included in the simulation, but are not shown. The top row shows the energy landscape (cost function) and a close-up of the region around the true hole location at $(x, y) = (20, 20)$ mm. The red markers indicate the measurement locations, and the cyan ring is located one half wavelength from the center of the hole. The bottom row shows the finite difference gradient field of the cost function and that computed from the shape gradient. The vectors all have unit magnitude and are intended to show direction only. The red circle indicates the outline of the 5 mm hole in the plate.

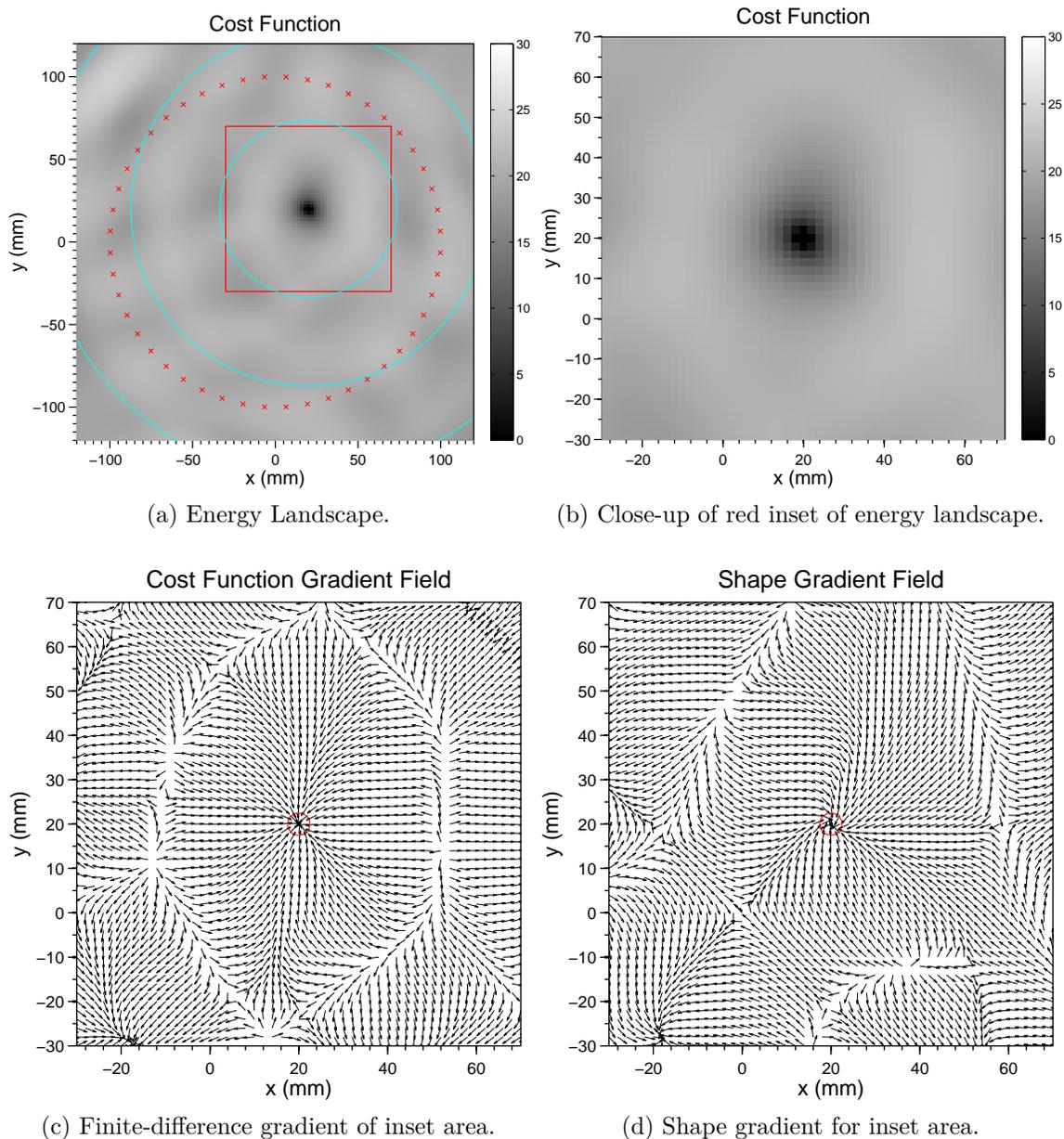


Figure 5.12: A numerical simulation of shape gradient reconstruction was conducted for S_0 waves at 50 kHz with 48 measurement points forming a ring about the origin of the scan area. Sources 1, 3, and 5 were included in the simulation, but are not shown. The top row shows the energy landscape (cost function) and a close-up of the region around the true hole location at $(x, y) = (20, 20)$ mm. The red markers indicate the measurement locations, and the cyan rings are spaced at half-wavelength intervals from the center of the hole. The bottom row shows the finite difference gradient field of the cost function and that computed from the shape gradient. The vectors all have unit magnitude and are intended to show direction only. The red circle indicates the outline of the 5 mm hole in the plate.

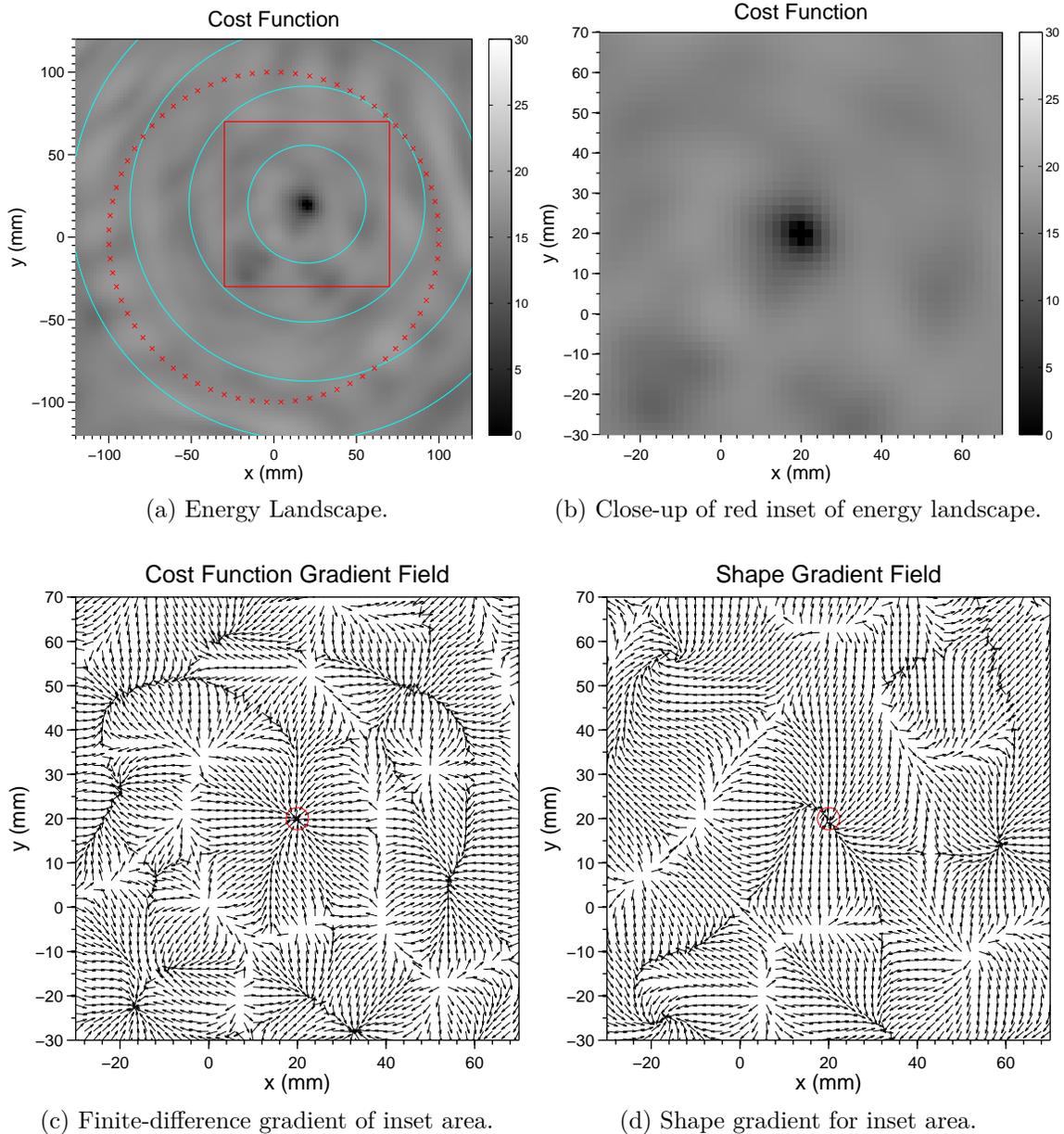


Figure 5.13: A numerical simulation of shape gradient reconstruction was conducted for S_0 waves at 75 kHz with 72 measurement points forming a ring about the origin of the scan area. Sources 1, 3, and 5 were included in the simulation, but are not shown. The top row shows the energy landscape (cost function) and a close-up of the region around the true hole location at $(x, y) = (20, 20)$ mm. The red markers indicate the measurement locations, and the cyan rings are spaced at half-wavelength intervals from the center of the hole. The top row shows the energy landscape (cost function) and a close-up of the region around the true hole location at $(x, y) = (20, 20)$ mm. The bottom row shows the finite difference gradient field of the cost function and that computed from the shape gradient. The vectors all have unit magnitude and are intended to show direction only. The red circle indicates the outline of the 5 mm hole in the plate.

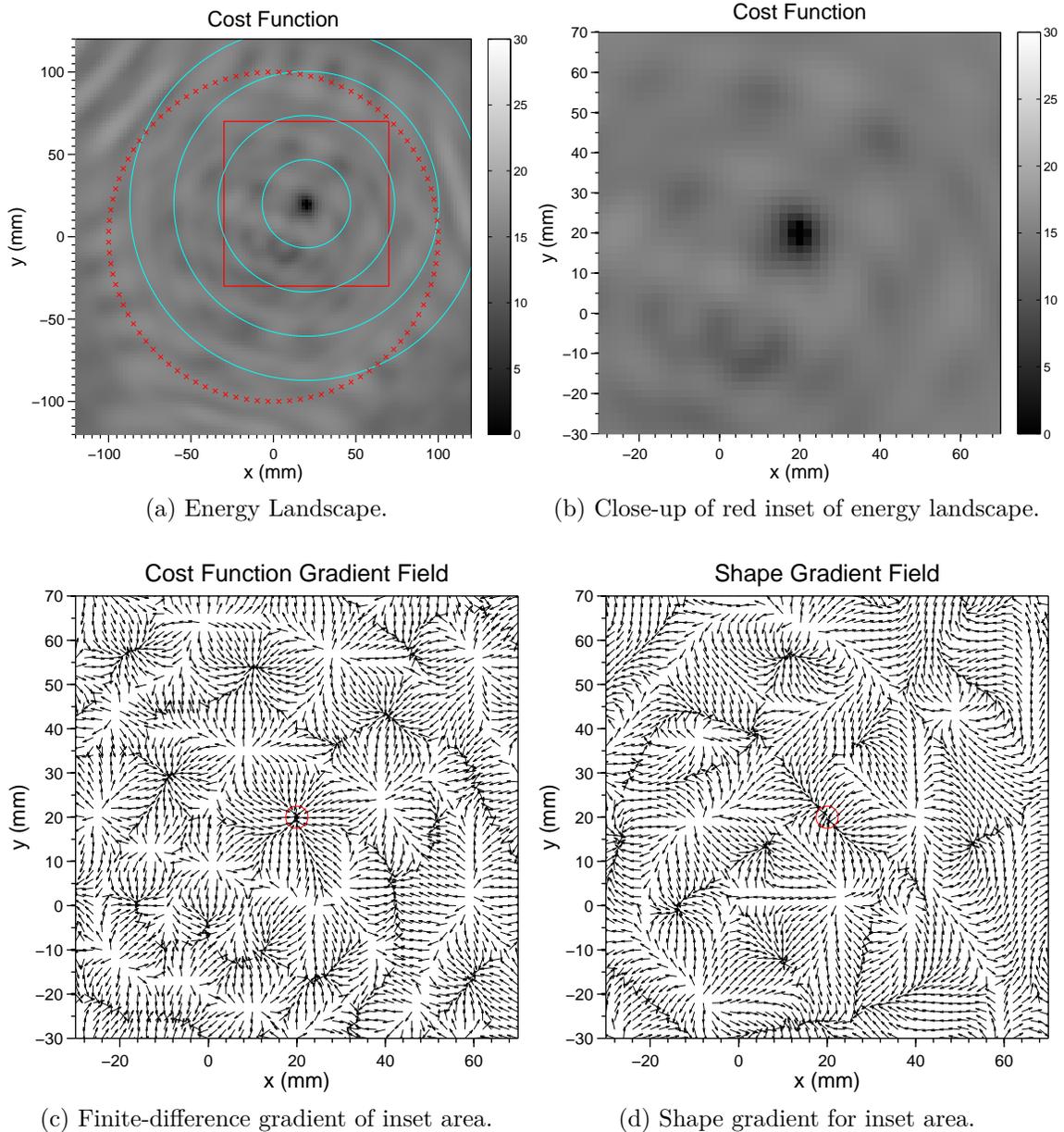


Figure 5.14: A numerical simulation of shape gradient reconstruction was conducted for S_0 waves at 100 kHz with 96 measurement points forming a ring about the origin of the scan area. Sources 1, 3, and 5 were included in the simulation, but are not shown. The top row shows the energy landscape (cost function) and a close-up of the region around the true hole location at $(x, y) = (20, 20)$ mm. The red markers indicate the measurement locations, and the cyan rings are spaced at half-wavelength intervals from the center of the hole. The top row shows the energy landscape (cost function) and a close-up of the region around the true hole location at $(x, y) = (20, 20)$ mm. The bottom row shows the finite difference gradient field of the cost function and that computed from the shape gradient. The vectors all have unit magnitude and are intended to show direction only. The red circle indicates the outline of the 5 mm hole in the plate.

PDEs and boundary conditions involved. Although the active surface reconstruction isn't compatible with the experimental data, other approaches could be used in its place. Since the shape being reconstructed is simple, it is possible to appeal to more traditional optimization strategies to find the unknown parameters. Two of these are outlined below.

5.6 Shape Recovery via Nonlinear Optimization

Optimization approaches such as nonlinear least squares (NLS) and simplex methods offer advantages under certain conditions. These are particularly useful when the shape can be represented by a small number of descriptive parameters. For the Lamb wave experiment only the hole's radius and its position in a 2D plane are needed—a total of three parameters. Although NLS and simplex method solutions are best suited to low numbers of unknown parameters, they have the distinct advantage of being amenable to the desired physical model. In contrast, the approximation used to obtain the shape gradient derived for the Lamb wave reconstruction is valid only for very large or very small values of ka .

In addition to the objective function, NLS ideally makes use of the corresponding Jacobian and Hessian matrices (first and second derivatives). This is by no means required, as solution schemes exist that employ numerical derivatives and/or slower rates of convergence [97]. Simplex methods are derivative-free, being designed to employ only the objective function. Both approaches are outlined briefly in this section.

5.6.1 Nonlinear Least Squares

Nonlinear least squares techniques are used to find a parameter vector $v \in \mathbb{R}^N$ from a set of M measurements, where it is understood that $M > N$. This solution is therefore applicable for shape reconstruction when the shape has a very simple form, such as a circle or sphere, which can be described using only three or four parameters (position

and radius). Arbitrary shapes are likely to require many descriptive parameters, and such reconstructions are better suited to the strengths of the adjoint approach, as detailed in Section 3.3.

The most general NLS solution is based on Newton's method and is described in Chapter 10 of Dennis and Schnaebel [97]. It is similar in form to the one-dimensional Newton's method solution for finding local extrema:

$$x_{n+1} = x_n - f'(x_n)/f''(x_n). \quad (5.26)$$

There are only three unknowns in the problem of reconstructing the scattering void, its radius and location in the (x, y) plane. But, there are many measurements available within the scanning area of the laser vibrometer. The problem is therefore overdetermined, and the measurements are used to find the best estimate of the three unknown parameters, expressed as the vector $v = [x, y, r]^T$. The goal is to find the v that minimizes the quantity $f(v) = \frac{1}{2}D(v)^T D(v)$, where $D(v)$ is known as the residual function. The m^{th} entry of $D(v)$ is written, using (3.28), as:

$$D_m(v) = u(x_m) - \hat{u}_m,$$

where there are M measurements and $D(v) \in \mathbb{R}^{M \times 1}$.

The first and second derivatives of $f(v)$ are required to form the solution using Newton's method. The first derivative of $f(v)$ is

$$\begin{aligned} \nabla f(v) &= \sum_{m=1}^M D_m(v) \cdot \nabla D_m(v) \\ &= J(v)^T D(v), \end{aligned} \quad (5.27)$$

where $J(v) \in \mathbb{R}^{M \times 3}$ is the Jacobian matrix with $J(v)_{mn} = \partial D_m(v)/\partial v_n$. Here, the m index denotes the m^{th} measurement and n indexes over the components of v . The rows of the matrix $J(v)$ are evaluated using known quantities and the current estimate

of v . The second-derivative of $f(v)$ is:

$$\begin{aligned}\nabla^2 f(v) &= \sum_{m=1}^M \left(\nabla D_m(v) \cdot \nabla D_m(v)^T + D_m(v) \cdot \nabla^2 D_m(v) \right) \\ &= J(v)^T J(v) + S(v),\end{aligned}\tag{5.28}$$

where

$$S(v) \equiv \sum_{m=1}^M D_m(v) \cdot \nabla^2 D_m(v).\tag{5.29}$$

The second derivative, or Hessian, of the residual $D_m(v)$ is $\nabla^2 D_m(v) \in \mathbb{R}^{3 \times 3}$ whose entries are given by:

$$\nabla^2 D_m(v)_{ij} = \frac{\partial^2 D_m(v)}{\partial v_i \partial v_j}.\tag{5.30}$$

It is evaluated using the same information used to evaluate the Jacobian matrix. Using the expressions for the first and second derivatives of the residual function, and by analogy with (5.26), Newton's method applied to the nonlinear least squares problem is written as:

$$v^c = v^p - \left[J(v^p)^T J(v^p) + S(v^p) \right]^{-1} J(v^p)^T D(v^p),\tag{5.31}$$

where v^c is the current estimate of the vector v computed from the previous estimate v^p , or from the initial guess. Equation (5.31) is iterated until the norm of the residual is below a specified threshold:

$$\|D(v)\| = \left[D(v)^T D(v) \right]^{1/2} < \epsilon_{\text{tol}}.$$

Once this condition is met, the iteration stops and estimates for x , y , and r are obtained.

While desirable, using the full Newton's method is not always practical because of the difficulty associated with obtaining closed-form expressions for the Hessian matrix and/or the expense in computing it. A simplified approach results from ignoring the higher-order derivative terms embedded in $S(v)$

$$v^c = v^p - \left[J(v^p)^T J(v^p) \right]^{-1} J(v^p)^T D(v^p).\tag{5.32}$$

This form is known as the Gauss-Newton method. As might be expected, the Gauss-Newton method converges more slowly than Newton’s method. It is also prone to taking bad step sizes. Two common modifications exist for controlling the step size. The first is a simple scaling of the Jacobian term

$$v^c = v^p - \lambda^p \left[J(v^p)^T J(v^p) \right]^{-1} J(v^p)^T D(v^p), \quad (5.33)$$

which is known as the damped Gauss-Newton method. The parameter λ is chosen to prohibit step sizes that don’t reduce the residual function. The value of λ is usually updated in some way that adapts to the local behavior of the cost function. This method of control is called a line search, because the Gauss-Newton steps will be in the right direction, but may over/undershoot.

Any of these versions of NLS is likely to work for the Lamb wave scattering problem. The harmonic expansion model provides a way to obtain the necessary derivatives. However, closed-form derivatives of Bessel functions are often written using recursion relationships [98], and the resulting digital implementation can sometimes be unstable because of the precision loss associated with subtracting very large or very small numbers. For this reason, the recommended approach for NLS is the Gauss-Newton method using finite-difference approximations for the first derivative terms. The smoothness of the cost function is well understood, being related to the scale of the wavelength, so the finite difference derivatives should not present any difficulties.

5.6.2 Simplex Methods

The NLS solution depends on knowledge of the derivatives of the cost function, obtained either through closed-form expressions or finite differences. Simplex methods, on the other hand, work only on the cost function itself. For an n -dimensional space, a simplex is a generalized notion of a polygon (more properly, a polytope) with $n + 1$

sides. For example, a 2-simplex is a triangle, and a 3-simplex is a tetrahedron. Optimization schemes based on simplices compute the objective function at each vertex of the simplex [99, 100]. For minimization problems, the vertex with the highest value is reflected through the centroid of the simplex and the process iterated. When visualized in two or three dimensions, the polytope resembles an amorphous blob gradually making its way to a local minimum. For this reason, the simplex method is sometimes called the amoeba method.

As with NLS there are many kinds of simplex methods, varying in their convergence properties and complexity [101]. Faster-converging schemes adapt the step size of the vertex reflection in ways similar to those used for controlling the line search for NLS. The resulting polytope can change size and shape dramatically during the course of the optimization. A simpler, but slower, approach is to assign fixed small edge lengths to the polytope. In two dimensions, the resulting optimization looks like a triangle flip-flopping down the cost surface toward the local minimum.

It was previously recommended that an NLS reconstruction use finite differences instead of closed-form derivatives. Since this is tantamount to computing only the cost function itself, there seems to be little advantage to favoring either NLS or simplex over the other. If computational efficiency were a concern, then the decision could be made based on convergence properties.

CHAPTER 6

DETECTION OF REFRACTIVE ANOMALIES IN SAS IMAGERY USING OPTICAL FLOW

Traveling boluses induced by breaking internal waves [102] have recently been observed in synthetic aperture sonar (SAS) imagery, where their appearance resembles natural seabed ripple patterns [40, 42]. The boluses cause refraction that concentrates the acoustic energy in a manner resulting in regions of increased intensity followed downrange by reduced intensity.

The presence of this effect in sidescan sonar imagery was discovered in 2012 in data collected by the Norwegian Defence Research Establishment (FFI). The same area of sea floor was surveyed multiple times, and the apparent sand ripples shifted in a manner that was inconsistent with known sediment transport behavior in the area. Example imagery showing the refractive effect of these boluses is given in Figure 1.2 of Chapter 1, repeated here as Figure 6.1. Even to a skilled interpreter, any one of these images taken by itself appears to show the kind of sand ripples which are common throughout the world's oceans. Rippled formations are known to occur at nearly all depths and size scales [103, 104, 105]. For comparison, Figure 6.2 shows a set of true seafloor ripples.

Given the similarity in appearance, it seems likely that refractive effects are unknowingly present in SAS imagery collected in environments that support the formation of internal wave induced boluses (IWBs). At first, distinguishing between true bathymetric features and IWBs may appear to be of purely scientific interest. There are, however, significant practical applications. If their presence can be detected, then those portions of a survey containing IWBs can be repeated as necessary to ensure that no objects on the sea floor fail to be imaged. This is particularly important for minehunting, where the potential cost of overlooking targets is exceedingly high.

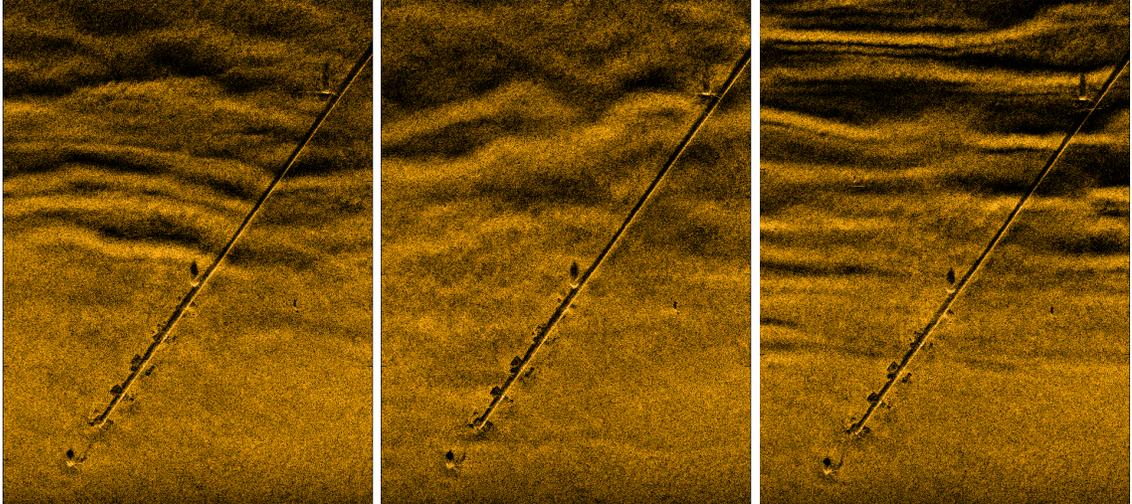


Figure 6.1: SAS images of the same scene were collected on three different passes on the same day, at times 07:19 (left), 08:18 (center), and 12:05 (right). The scene size is 80 m x 100 m and shows a pipeline with anchors on the seabed. The data were collected by FFT's HUGIN AUV carrying an interferometric SAS during the ARISE12 trials on board the NATO research vessel *Alliance*, outside Elba island, Italy.

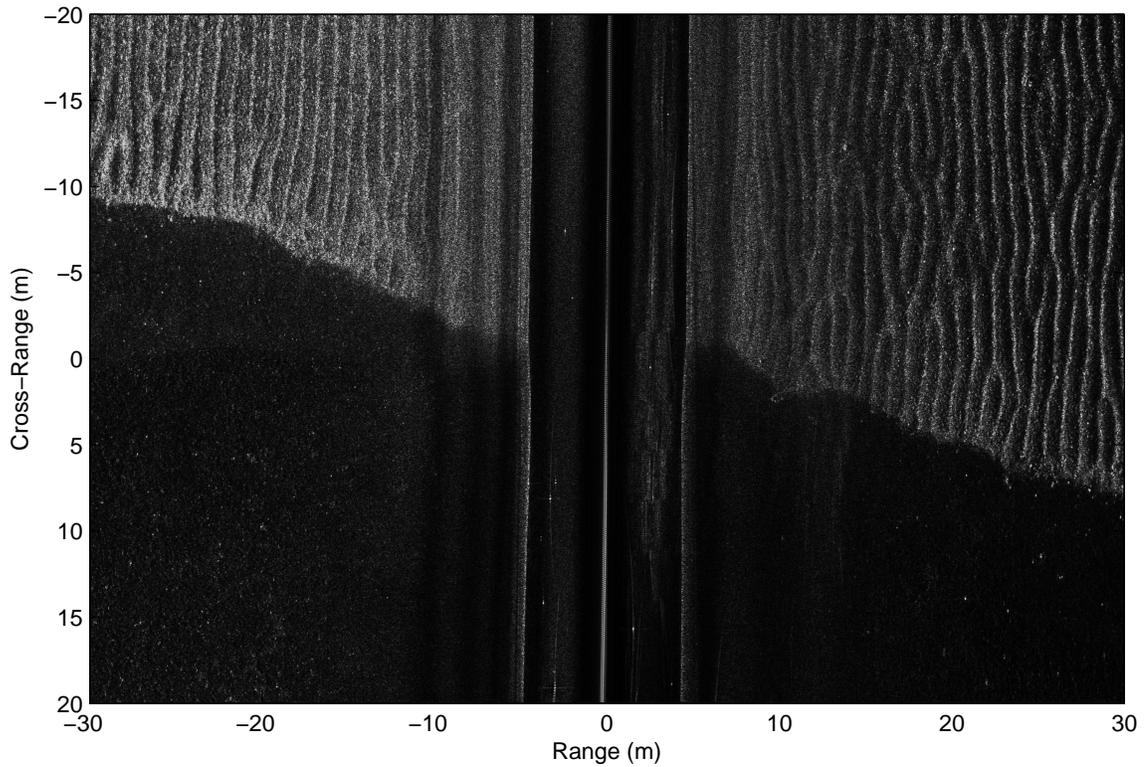


Figure 6.2: Sand ripples observed in the Gulf of Mexico by the Office of Naval Research SAS12 synthetic aperture sonar.

Detecting IWBs is obviously a time consuming process if repeat surveys are required. The motivation therefore exists to find detection methods that can be applied to single-pass imagery. Several options are outlined by Hansen et al. [42], and the expected performance of each is discussed. This chapter builds on that work, extending one of the proposed methods by using PDE-based optical flow to visualize apparent motion of the IWB. Optical flow is a well-established means of estimating frame-to-frame motion of video image content, and it may be used for single-pass detection and motion tracking of transient refractive effects in SAS imagery such as those induced by IWBs. Another goal of this work is to demonstrate the utility of variational methods for SAS-related applications by linking free parameters (or “knobs”) used to control the behavior of the algorithm to known physical and statistical properties of the imagery.

6.1 Breaking Solitary Waves

Internal waves are supported in ocean waters where a steep density gradient exists. This layering might be due to differences of temperature, or salinity, or both. The internal waves propagate along the density interface, or pycnocline, either singly or in groups. Isolated internal waves are called solitary waves or solitons. The fact that they are common in the ocean has only been firmly established in the past two decades. Solitons are governed by nonlinear effects, representing a balance between (1) a tendency for wave speed to increase with amplitude, causing steepening and shock formation, and (2) dispersion that tends to broaden wavefronts [106]. The resulting wave possesses a stable shape.

The specific feature of solitons concerning this research is their interaction with sloping sea floors. It has been shown numerically [107] and experimentally [108] that solitons traveling up-slope can break, separating a bolus of dense water that continues to travel along the sea floor before eventually dissipating. Figures 6.3 and 6.4 show

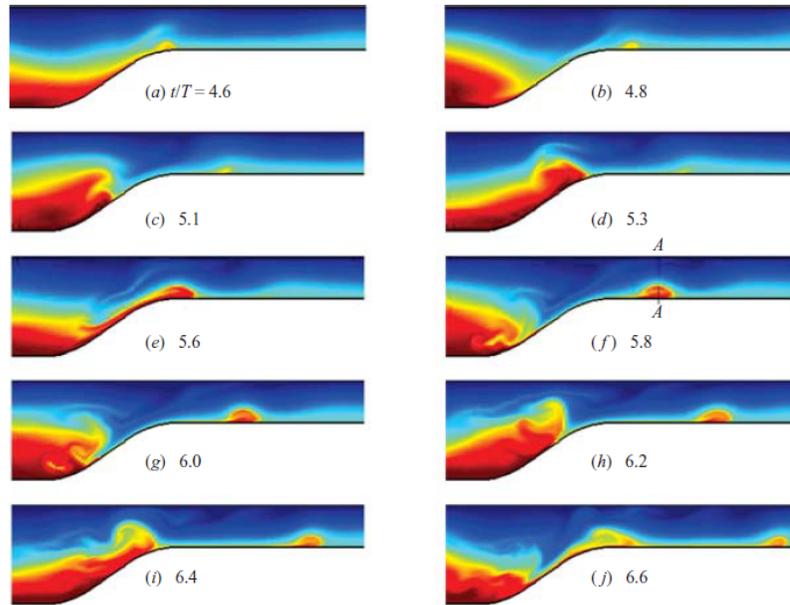


Figure 6.3: Numerical simulation of the density field associated with a breaking soliton and the resulting bolus. (Figure from Venayagamoorthy and Fringer [107].)

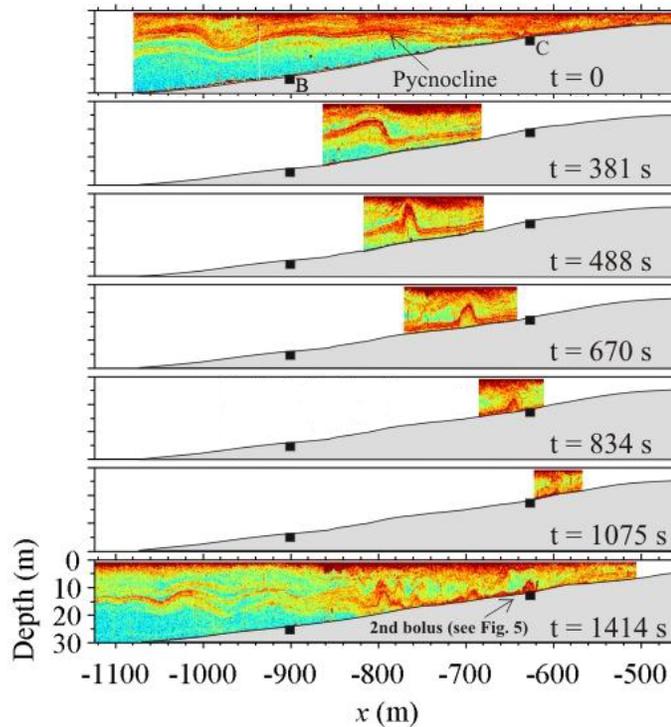


Figure 6.4: Acoustic soundings were used to observe soliton breaking off of Ile-aux-Lièvres Island in the St. Lawrence Estuary in 2004. (Figure from Bourgault et al. [108].)

examples from the literature. Lyons et al., point out that the size and speed of these boluses is consistent with the features visible in the 2012 FFI SAS imagery [40]. It is therefore the present consensus of the SAS community that internal wave induced boluses are responsible for the transient refractive effects observed in the imagery.

6.2 Optical Flow

The Lagrangian view of fluid mechanics may be used to motivate the derivation of optical flow methods. In an analogy with fluid flow, consider a small patch of image of fixed area whose intensity is represented as $I(x(t), t)$, where the position vector is $x(t) = [x_1(t), x_2(t)]^T$. The patch is able to translate in space and its intensity may vary. Taking the total derivative with respect to time gives

$$\frac{\partial I}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial I}{\partial x_2} \frac{dx_2}{dt} + \frac{\partial I}{\partial t} = \left(\frac{\partial}{\partial t} + v \cdot \nabla \right) I = \frac{DI}{Dt}, \quad (6.1)$$

where $v = [v_1, v_2]^T = [dx_1/dt, dx_2/dt]^T$ is the velocity vector, and the shorthand at the end of (6.1) is known as the material, or substantial, derivative [109]. The material derivative gives the instantaneous change in I at any given position and time.

When the intensity of each image patch is constrained to be unchanging as it moves (a condition known as brightness constancy or the optical flow condition), the resulting PDE becomes $DI/Dt = 0$, or

$$\frac{\partial I}{\partial t} = -v \cdot \nabla I. \quad (6.2)$$

This is a scalar PDE stating that the time rate of intensity change equals the component of the intensity gradient in the direction of motion. By itself, this equation cannot be solved for the velocity field. It only captures the component of the optical flow in the direction of ∇I ; that is, the flow normal to contours of constant intensity. This ambiguity of motion is known as the aperture problem, and there are a number

of approaches for resolving it [110]. One possibility is to impose a smoothness constraint on v . The literature contains many such constraints, tailored to a variety of conditions.

The method of Horn and Schunck [111] was chosen for this investigation of optical flow for detecting the presence of transient refractive effects. Here, brightness constancy is combined with a smoothing term to obtain an energy functional that is minimized by the optimal optical flow field:

$$\begin{aligned}
 E &= (1 - \lambda) \int_{\Omega} \left(v \cdot \nabla I + \frac{\partial I}{\partial t} \right)^2 dx + \lambda \sum_{j=1}^2 \int_{\Omega} \|\nabla v_j\|^2 dx \\
 &= \int_{\Omega} \mathcal{L}(v, v', x, t) dx,
 \end{aligned}
 \tag{6.3}$$

where Ω is the domain of integration (i.e., the image), v' represents the possible derivatives of v with respect to its arguments (e.g., $\partial v / \partial x_1$ and $\partial v / \partial x_2$), and $\lambda \in [0, 1]$ is a user-chosen parameter whose purpose is to balance the relative contributions of the data fidelity ($v \cdot \nabla I + \partial I / \partial t$) and smoothness ($\|\nabla v_j\|$) terms comprised by the Lagrangian \mathcal{L} . Horn and Schunck use a slightly different weighting α , which is equivalent to $\lambda / (1 - \lambda)$. They suggest that the square of this term should be roughly equal to the expected noise level of the image gradient. It is shown below that this quantity can be estimated using the well-established statistical properties of coherent scattering from rough surfaces.

The calculus of variations is used to find the function $v(x, t)$ that minimizes the energy functional E . This is accomplished by solving the Euler-Lagrange equations [112, 113] derived from \mathcal{L} according to the following condition necessary for minimizing E :

$$\frac{\partial \mathcal{L}}{\partial f_i} - \sum_j \frac{\partial}{\partial \xi_j} \left(\frac{\partial \mathcal{L}}{\partial f'_{i,j}} \right) = 0,
 \tag{6.4}$$

where the subscript i indicates the components of the vector function f , ξ_j is the j^{th} argument of f , and $f'_{i,j} \equiv \partial f_i / \partial \xi_j$. For this problem, v is identified with f . Carrying

out the calculations, a pair of coupled equations is obtained:

$$\begin{aligned} (1 - \lambda) \left(v \cdot \nabla I + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial x} - \lambda \nabla^2 v_1 &= 0 \\ (1 - \lambda) \left(v \cdot \nabla I + \frac{\partial I}{\partial t} \right) \frac{\partial I}{\partial y} - \lambda \nabla^2 v_2 &= 0. \end{aligned} \tag{6.5}$$

There are a number of suitable numerical approaches for solving the Euler-Lagrange equations for v . A simple one is gradient descent, in which an artificial dependence on a time marching parameter τ is assigned to v . Re-deriving the Euler-Lagrange equations to include this dependence results in the same pair of equations above with $[\partial v_1/\partial\tau, \partial v_2/\partial\tau]^T$ on the right-hand side instead of $[0, 0]^T$. Starting from an initial estimate v_0 and the measured image pair, the solution is iterated while $[\partial v_1/\partial\tau, \partial v_2/\partial\tau]^T$ approaches $[0, 0]^T$. In practice an appropriate stopping criteria is specified to keep the number of iterations reasonable; for example, the norm of $[\partial v_1/\partial\tau, \partial v_2/\partial\tau]^T$ falling below a given threshold that is close to zero. Alternatively, the largest dimension of the image may be used to set a fixed number of iterations.

6.3 Image Frame Generation

Optical flow computations are typically applied to sequential video frames. For sonar, repeat passes are the obvious way to obtain such a sequence, but this is highly undesirable. The time between passes is often long, and re-imaging severely reduces the area coverage rate of the survey. The goal of this work is to find a reliable approach for detecting transient refractive effects within a single image.

There are a number of ways in which a single SAS data collection might be divided into subsets that can be processed into sequential images. Some of these are discussed by Hansen et al. [42]. The approach chosen for this research begins with a complex-valued stripmap SAS image, as opposed to the raw sonar data it was derived from. The image is Fourier transformed into the two-dimensional (k_x, k_y) wavenumber domain, allowing the data to be easily filtered according to aspect, θ_{view} , and integration

angle, θ_{int} , by applying the appropriate mask to the spectrum. The region of interest is defined by $\mathcal{M}(k_x, k_y) = 1$ for

$$\left. \begin{array}{l} k_x = 2k \sin \theta \\ k_y = 2k \cos \theta \end{array} \right\} \theta \in \left[\theta_{\text{view}} - \frac{\theta_{\text{int}}}{2}, \theta_{\text{view}} + \frac{\theta_{\text{int}}}{2} \right], \quad (6.6)$$

and zero otherwise. Additional masking and/or weighting to control the shape of the point scatterer response could be built into \mathcal{M} . Applying an inverse Fourier transform to the masked spectrum provides a single image frame. An image sequence is constructed by varying θ_{view} . Beams of various widths and degree of overlap can be extracted, and the resulting set of images used as input to the optical flow. For example, a single image might be decomposed into a pair of images where one corresponds to the forward-looking half of the sonar beam and the other to the aft-looking half.

The performance of the optical flow depends on θ_{view} and θ_{int} . Positive/negative values of θ_{view} map to beams steered fore/aft, and there is a corresponding time delay between the images formed from these beams. Since the images are separated by angle, the delay between images depends on range according to $\Delta t = v_s^{-1}y [\tan(\theta_{\text{fore}}) - \tan(\theta_{\text{aft}})]$, where v_s is the forward speed of the sensor, y is range, and θ_{fore} and θ_{aft} are the values of θ_{view} for the pair of images being compared. Consequently, any image-to-image displacement vector measured by optical flow can be converted to a velocity by dividing it by the appropriate range-dependent Δt .

There is a trade associated with the selection of integration angle for the generation of frames. Larger angles are favored for imaging because they give finer cross-range resolution. When the object of interest is moving, smaller integration angles are often more suitable because the motion may be negligible or nearly constant over the time required to collect the data. This effect has been used for detecting moving objects in SAR imagery. Fine resolution is unnecessary in the present context, so θ_{int} can be small to improve temporal localization. Table 6.1 summarizes the parameters available to the user when applying Horn and Schunck optical flow to SAS imagery.

Table 6.1: Parameters available for controlling optical flow.

Parameter	Comments/Effect
α	Mean of speckle gradient
Image oversampling factor	Apparent smoothing of image
Aspect angle, θ_{view}	Time between image frames
Integration angle, θ_{int}	Image acquisition time and cross-range resolution

6.4 Experimental Approach

All optical flow schemes are affected by noise in the image, resolution, and the amount of change between consecutive images, which are typically fixed by the sensor and imaging conditions. The SAS application is somewhat unique in that it offers the flexibility to vary all of these. The parameters are coupled, but the user has tremendous latitude nevertheless.

6.4.1 Weighting the Smoothness Constraint

A key property of SAS is that an image can be modeled as an underlying reflectivity function modulated by the speckle caused by the coherent rough scattering from the sea floor. Speckle is not noise in the traditional sense because it is an inherent property of coherent imaging systems. Furthermore, speckle is “deterministically random” in that it is caused by the random distribution of small-scale scatterers throughout the scene, yet that realization is the same each time the scene is imaged. This feature is widely exploited for change detection and interferometric terrain height estimation [3, 114], as well as motion estimation for SAS [5]. The speckle realizations become decorrelated as two views of the same scene are separated in angle. This effect limits the baseline that can be used for interferometry [115]. It is used to advantage, however, in multilook processing where the image spectrum is divided to make a set of images with lower resolution, but independent speckle realizations [116, 3]. The

images are averaged, thus reducing the speckle contribution and revealing details not otherwise visible. Multilooking is commonly done for remote-sensing SAR, as these collections typically have ample spatial bandwidth to work with. It is less commonly used for SAS because the nature of sound propagation in the ocean makes it expensive to construct sensors capable of high sampling rates in the along-track direction. Traditional multilooking tiles the wavenumber spectrum in both k_x and k_y . Since the image frames for optical flow are generated by excising a specifically chosen subset of the available beamwidth, any multilooking is applied only in the k_y dimension.

Speckle can be thought of as a modulation varying between 0 and 1 that imparts a grainy texture to the image. For high-quality SAS imagery, the speckle is far more pronounced than the noise. Assuming this to be the case, it is reasonable to follow the guidance of Horn and Schunck and set the weighting parameter α^2 equal to the average value of the squared magnitude of the speckle (in lieu of noise) gradient. The granularity of speckle is inversely proportional to the spectral support of the image. Therefore, the average size of speckle is equal to the resolution of the image, regardless of the sampling rate, and even regardless of the quality of image focus. For this reason, the apparent gradient of the speckle depends on the sampling rate of the image. Oversampling causes the speckle to appear smoother when its gradient is computed numerically.

The mean value of the speckle gradient cannot generally be computed directly from the image, although a smooth region of sea floor may provide a reasonable value. Regardless, it can be conveniently estimated from a speckle realization created separately from the image, yet possessing the same spectral support and weighting. This is done by populating the wavenumber spectrum with complex phasors, whose phases are uniformly distributed, and then multiplying it by the mask \mathcal{M} defined earlier for the purpose of creating image frames. The steps are outlined as follows:

1. Compute the two-dimensional Fourier transform of the complex-valued SAS image
2. Threshold the spectrum to create a binary mask indicating the region of support in the Fourier domain
3. Populate the region of support with unit-amplitude phasors with uniformly-distributed phase
4. Compute the inverse Fourier transform to obtain a speckle realization
5. Compute the expected value $\langle I_x^2 + I_y^2 \rangle$ as an estimate of α

Alternatively, the excellent book by Goodman [117] contains an extensive closed-form analysis of speckle statistics that could be used to compute α .

6.4.2 Image Resolution

Fine-resolution SAS imagery has resolution on the order of 2-4 cm, while the refractive distortions present in the image are typically on the scale of meters. Retaining the best possible resolution is unnecessary for detecting transient refractive effects, so the SAS image therefore has excess spatial bandwidth that can be used to reduce speckle via multilooking. Multilooking carries the added benefit of reducing the pixel count in the image, greatly accelerating the optical flow computation. This reduction is advantageous because SAS imagery often contains on the order of 10-100 million pixels, and the optical flow can require thousands of iterations resulting in impractically-long computation times. The spatial filtering and multilooking applied to the SAS data in this experiment resulted in images with approximately 5% of their original pixel count, a far more manageable size.

6.4.3 Spectral and Temporal Overlap

In addition to the integration angle (resolution), the angle between frames (time separation) must be decided. These angles can be such that there may or may not be

overlap in the spectral support between two or more consecutive frames. It is possible to construct a sequence of images with arbitrarily small time separation. There are drawbacks to making the time separation too large or too small. If the inter-frame time separation is too small, then slow-moving features in the image may not change enough to be reliably detected. If it is too large, then the scene content may change drastically resulting in a meaningless optical flow field.

Regardless of the integration angle, a key consideration is the amount of spectral overlap between consecutive frames. If there is no overlap, then the speckle realizations will be completely uncorrelated. Otherwise, the speckle will exhibit some degree of correlation. Note that the correlation of the speckle is in addition to any correlation due to the underlying seafloor reflectivity, for example, from specular or point-like reflectors. These facts make it difficult to predict the correlation between frames, but it is clear that as the overlap of the frames' spectral regions of support decreases, the optical flow field can be expected to exhibit some amount of fine scale randomness [118]. This randomness can be masked or altered by the smoothing term in Equation (6.3) and by the number of iterations used to compute the optical flow field.

6.5 Experimental Results

The optical flow estimation was applied to data collected by FFI's HUGIN autonomous underwater vehicle carrying a HISAS wideband interferometric SAS. Two collections were chosen for comparison. Both contain scenes of a smooth sea floor, where one image is free of refractive anomalies while the other is known to contain them. The presence of the anomalies was confirmed using imagery from repeat passes. (Recall Figure 6.1, and see the article by Hansen et al. [42].)

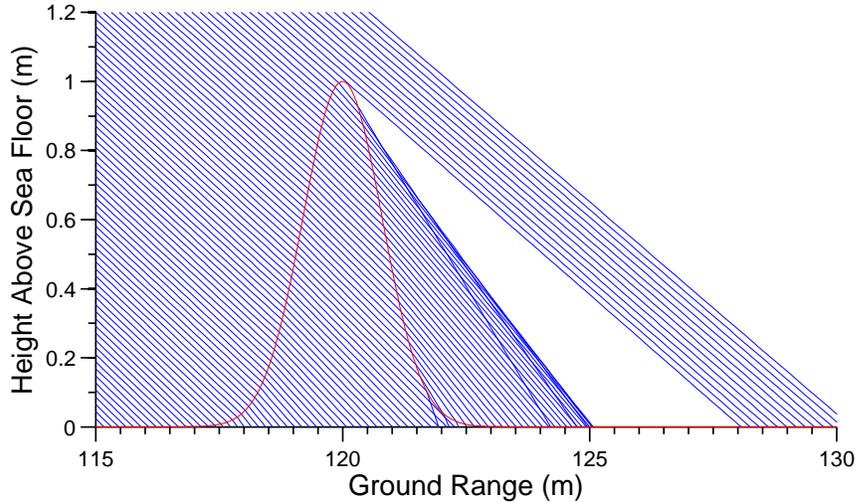


Figure 6.5: Rays were traced through a hypothetical internal wave bolus using a geometry similar to the collection used for this study. Note the concentration of acoustic energy followed downrange by a region that would appear as a shadow in an image.

6.5.1 Raytracing

To better understand the behavior of optical flow in the context of SAS, the refraction of an IWB was modeled by tracing rays through a local anomaly in the sound speed. The choice of this shape is somewhat arbitrary, as ground truth is not available. A parabola was used in [42], and a Gaussian shape is used here, giving similar results. An example raytrace is shown in Figure 6.5, where the anomaly is centered at a ground range of 120 m and extends approximately 2 m downrange and uprange. The most obvious features visible in the raytrace are the concentration of acoustic energy just downrange of the IWB and the depletion of acoustic energy causing an apparent shadow to follow it. It is this combined highlight/shadow that causes the refraction associated with the IWB to strongly resemble the appearance of sand ripples.

The Gaussian shape used in Figure 6.5 contains a feature not present in the parabolic model. This is the secondary focusing at the downrange “foot” of the IWB, located at a range of about 122 m in the figure. The raytraced profile was repeated to create a simple reflectivity map. Speckle with the appropriate properties was then

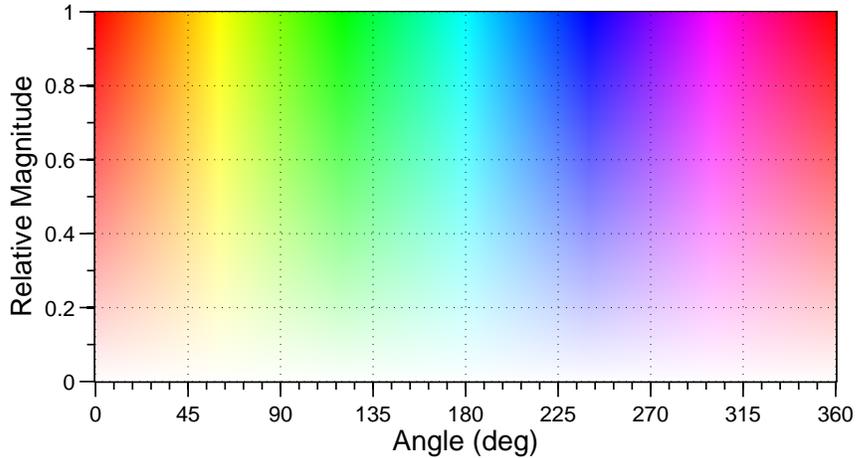


Figure 6.6: Legend for HSV optical flow representation. The color (hue) represents the angle of the optical flow field, while the magnitude is given by saturation. Angles of 90° and 270° correspond to the downrange and uprange directions, respectively.

included to create a simulated SAS image. These images are shown in the top row of Figure 6.7. Which model is used for the IWB is largely immaterial for the purpose at hand. However, it is interesting to note that some of the refractive features visible in the right panel of Figure 6.1 exhibit the paired highlights predicted by the Gaussian model. It is also important to point out that the image artifacts appear downrange of the actual location of the IWB.

Optical flow estimation was applied to a sequence of images created by stepping the Gaussian shape downrange. An independent speckle realization was applied to each frame in keeping with the choices of θ_{view} and θ_{int} used in the next section for the sea test data. The results are presented in Figure 6.7. The left and right columns show the optical flow results for the raytraced image without and with speckle, respectively. The middle row shows the magnitude of the optical flow. The true speed of the simulated IWB was 2 cm/s. The bottom row plots the direction of the optical flow vector field as well as its normalized magnitude using an HSV (hue-saturation-value) representation. The hue represents the angle of the flow vector field, while the saturation indicates the magnitude. Since the magnitude is provided separately,

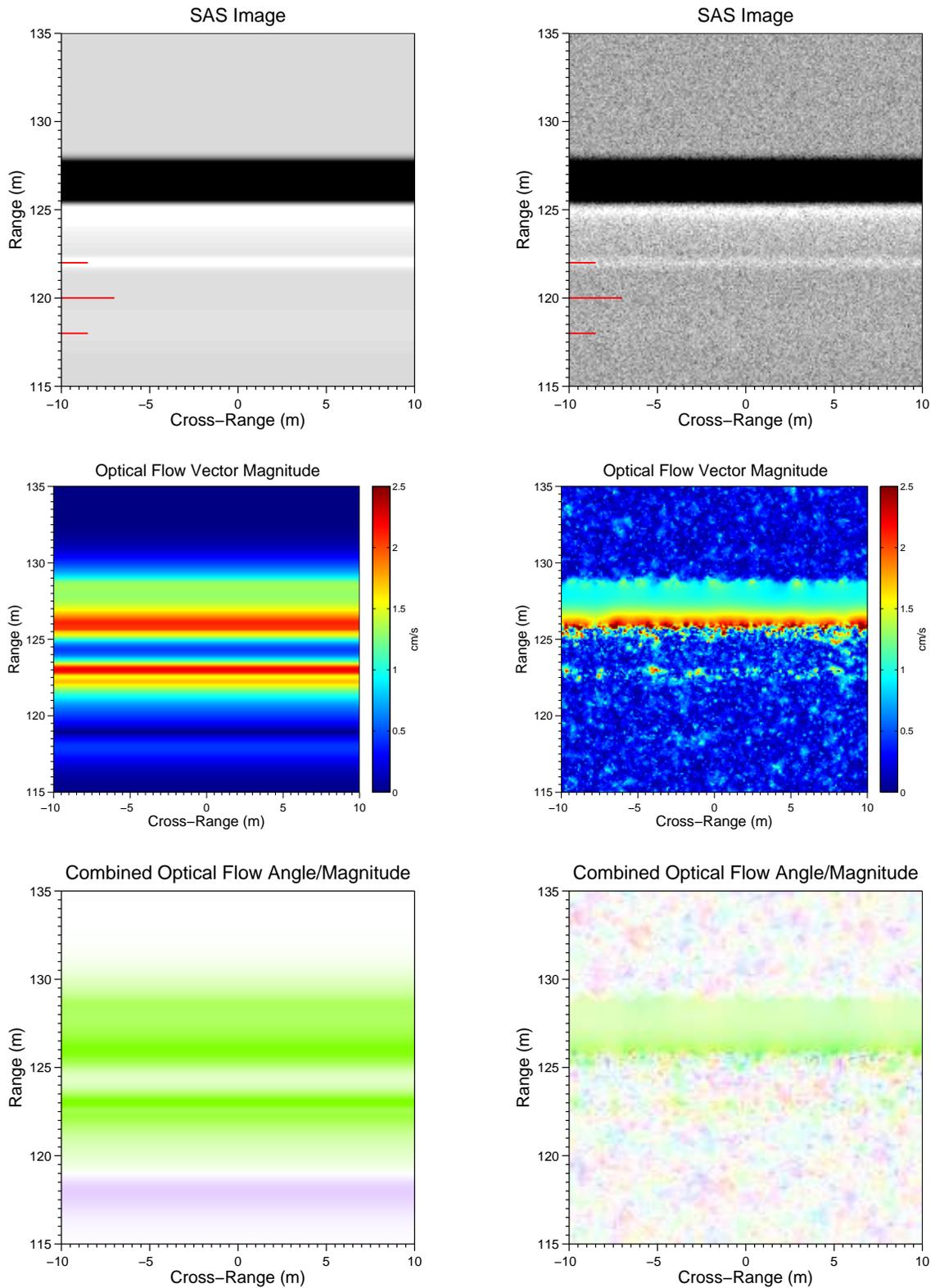


Figure 6.7: Comparison of optical flow for images without (left column) and with (right column) speckle effects. The simulated SAS image was generated from the raytrace shown in Figure 6.5. The red line segments indicate the ground ranges of the center and edges of the IWB.

the saturation is normalized to the maximum flow magnitude to enhance the angle visualization. The mapping between flow angle and HSV color is given by the legend in Figure 6.6.

The speckle-free optical flow in Figure 6.7 captures the details of the time-varying intensity due to the movement of the refractive bolus. The addition of speckle removes much of this detail, leaving a single dominant artifact. The optical flow appears to lose details whose downrange extent is small, or at least comparable to the characteristic dimension of the speckle. The highlights caused by the refractive focusing fall into this category, and are difficult to discern in the bottom right plot of Figure 6.7. The motion of the shadow is the dominant feature captured by the optical flow. This is because the shadow is a large area of uniform (albeit low) intensity and because there is little speckle in this region.

6.5.2 Application to Sea Test Data

The optical flow was applied to two pairs of images collected by FFI using the SAS carried by the HUGIN autonomous underwater vehicle. The first image pair, Figure 6.8, contains a single strong IWB. The second set, Figure 6.9 shows the optical flow applied to an image that is completely filled with refractive effects. Both examples contain a ‘control’ image that was taken at the same range but is free of refractive effects. The images in Figure 6.8 are of the same area of sea floor. The areas imaged in Figure 6.9 are near one another, but not identical.

As with the raytraced simulation, the top row of Figures 6.8 and 6.9 shows one frame from the image pair used to compute the optical flow. The middle row shows the magnitude of the optical flow vector field, and the bottom row shows its angle. The effective integration angle is 1° and the frames correspond to view angles of $\pm 5^\circ$. It is important to stress that the optical flow field is not a direct measurement of the motion of the transient refractive anomaly. Even a totally quiescent environment is expected to exhibit some amount of optical flow because of differences in the speckle

realizations. Furthermore, the multispect image frames will cause shadows to be cast in slightly different directions [119]. The observed optical flow field is a product of the following contributions:

1. Differences in speckle between image frames
2. True displacement of the IWB
3. Evolving shape of the IWB
4. Apparent motion due to frame-to-frame variation in θ_{view}
5. Motion of the IWB during the time needed to form the synthetic aperture

These factors should be taken into account if it is desired to go beyond detection and use the SAS imagery derive quantitative information about the true IWB.

The imagery examined in Figure 6.8 is helpful for characterizing the optical flow performance because the IWB is clearly isolated within the image. One therefore expects the majority of the optical flow field to behave randomly, as for the IWB-free image. The large IWB spans the width of the image, but only about 40% of it is clearly visible in the optical flow. If the constituent image pair is compared carefully (not shown), the downrange limb of the IWB is seen to move appreciably more than the rest of the feature. It is not clear that the near-range part of the IWB is moving at all in the image sequence used. There is also a slight narrowing of the image artifact at its closer ranges that would tend to mask what little motion may be present.

The optical flow angle indicates that the IWB is moving uprange. However, the collection geometry and local bathymetry indicate that the IWB is actually moving in the downrange direction. This discrepancy is believed to be attributable to the optical flow being dominated by apparent motion due to changing the view angle, as opposed to the true motion of the IWB. The fact that the IWB is at an oblique angle

relative to the along-track axis of the image complicates the refraction. It should be possible to model this by extending the raytracing approach described above.

In contrast with the solitary IWB, Figure 6.9 shows the result of applying optical flow to a sizable region of a SAS image that is completely filled with refractive effects. The image containing the refractive effects was taken from the upper-left corner of the right-hand image in Figure 6.1. No objects are visible in the excised region, but it can be inferred from the remainder of the image that the scene is well-focused. The refraction-free comparison image contains some man-made objects large enough to cast shadows, but otherwise shows a uniform sea floor. The shadows are seen to induce some very localized motion, as they shift in response to the view angle.

The refractive boluses shown in the figure are undulating, but roughly parallel to the cross-range axis, and are known to be moving quite slowly in the downrange direction. It is clear that the overall optical flow is parallel to the range axis, but it is not uniform. Rather, it is characterized by distinct regions moving both uprange and downrange. The reason for the sign change is not yet understood. It is possible that view-to-view changes in refraction are responsible, as is believed to be the case in Figure 6.8. Another prominent visible effect are the unstructured regions appearing as nearly horizontal stripes at ranges of 110 m, 120 m, and 128 m. These locations coincide with the highlights in the SAS image. They appear to match the phenomenon visible in the raytraced imagery (Figures 6.5 and 6.7) at the range corresponding to the most intense ray focusing. This confusion of the optical flow field seems reasonable, given the crossing rays visible in Figure 6.5.

6.6 Discussion

The optical flow behavior is impossible to explain completely without accurate knowledge of the propagation environment. Nevertheless, some qualitative expectations can be stated: The angular-temporal nature of the multiaspect filtering used to create the

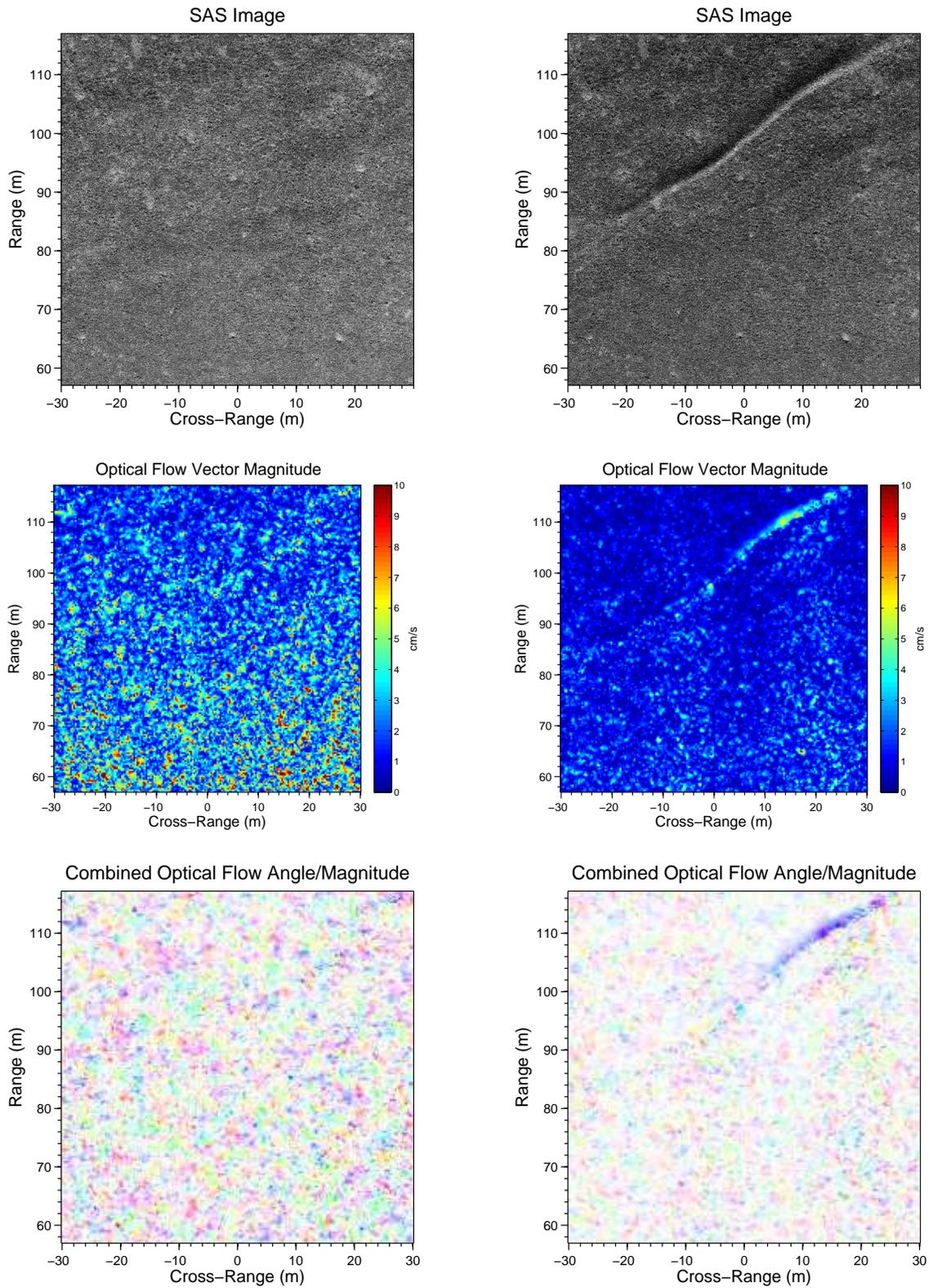


Figure 6.8: Comparison of optical flow for images without (left column) and with (right column) refractive effects. The top row shows the underlying SAS imagery, while the middle and bottom rows show the magnitude and angle of the optical flow vector field.

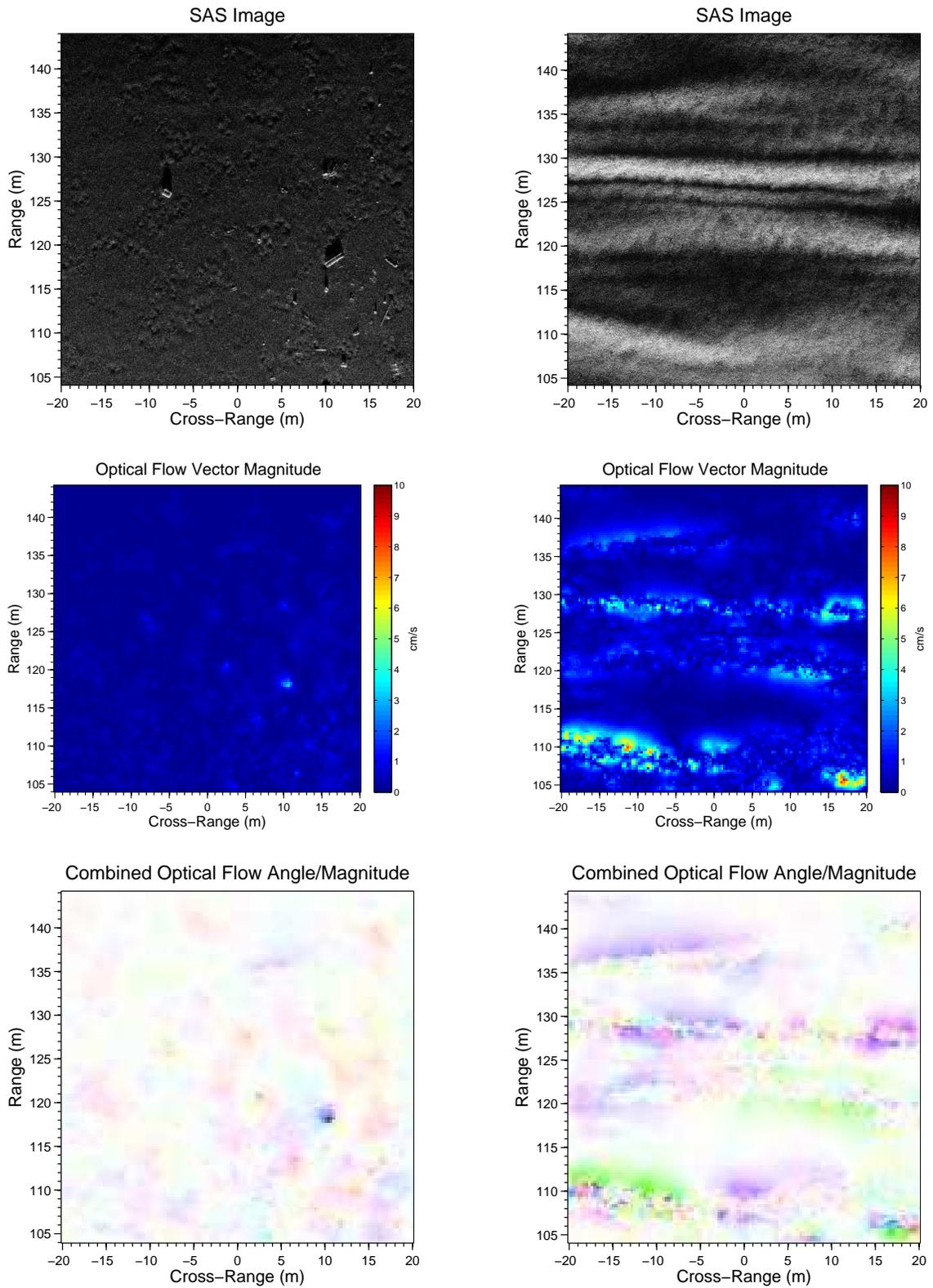


Figure 6.9: Comparison of optical flow for images without (left column) and with (right column) refractive effects. The top row shows the underlying SAS imagery, while the middle and bottom rows show the magnitude and angle of the optical flow vector field.

image frames should produce two kinds of effects: First, any changes in the speckle pattern should result in a randomization of the optical flow field. Secondly, when objects are present in the scene, their highlights and shadows will shift primarily in the cross-range direction according to the change in θ_{view} . The corresponding optical flow vectors should point in the cross-range direction as well (angles near 0° and 180°).

Assuming the speckle is uncorrelated among the two images being compared, it should be possible to predict an upper bound on its optical flow. Since the characteristic dimension of the speckle is the same as the image resolution, the optical flow displacement should not exceed this distance, regardless of the temporal separation or amount of frame-to-frame decorrelation. For the parameters used to visualize the FFI data, the upper bound on speckle-driven optical flow is approximately 6 cm/s, and the mean value is of course substantially lower. This prediction is supported by the observations in Figures 6.8 and 6.9.

The characteristics of speckle are well-understood, suggesting an approach for automating the detection of refractive anomalies. For a given scene, a sliding window can be used to estimate the local magnitude and direction of the optical flow map. Those areas deviating significantly from the behavior associated with pure speckle would be marked as being possibly influenced by refraction. Strictly speaking, speckle locally violates the brightness constancy assumption used to derive optical flow, behaving as a continuum of sources and sinks of intensity such that $DI/Dt \neq 0$. However, brightness constancy is obeyed on average.

CHAPTER 7

SUMMARY AND CONCLUSIONS

This research was motivated by a desire to bring ideas from PDE-based image processing into the sonar, radar, and nondestructive testing communities. Although isolated examples are cited in Chapter 1, there has not yet been widespread exposure within these fields. Image reconstruction via aperture synthesis is an important application in all of these disciplines, so it was chosen as the context in which variational techniques are discussed. Similarities and differences among a number of imaging modalities presented in Chapter 2 include a high-level mathematical description of image reconstruction via the adjoint to the model representing the observed data. The adjoint operator was shown to be closely linked to the notion of backprojection, and most, if not all, of the major synthetic aperture reconstruction algorithms can be related back to this framework.

The first of two main thrusts of this research is the idea of the adjoint active surface whose cost function is linked directly to the physics governing the wave propagation and scattering. The primary distinction between this approach and traditional imaging is that the purpose of the active surface is to segment out a scatterer of interest from within a larger volume. While the entire volume is not imaged, a number of significant benefits accrue to offset this loss. All degrees of freedom available to the reconstruction are placed on the scattering surface of interest, not on the surrounding volume. The reconstruction is also effectively decoupled from the specific solution to the wave propagation. For example, the shape gradient needed for reconstruction is written in terms of the field values on the active surface. These may be computed in any way that is convenient to the user and well-matched to the problem. Since numerical techniques are appropriate to use, reconstruction in complicated environments becomes a more approachable prospect. An interesting feature of the adjoint active

surface reconstruction is that it appears to be well-suited to low and mid-frequency scattering, where many reconstructions are best suited to the high-frequency (short wavelength) regime.

Active surfaces were developed in the context of the Helmholtz equation for the cases of (1) opaque scatterers (i.e., with no interior field) embedded in free space, and (2) penetrable scatterers in a volume that may be bounded. The latter was demonstrated numerically using closed-form solutions based on spherical harmonics. The former case was chosen as the basis for a laboratory experiment using Lamb waves in an aluminum plate. The experiment was carefully designed so as to produce only the S_0 and A_0 modes in the plate. Since the modes travel with different speeds, it is possible to spatially filter the data to eliminate the unwanted mode. The shape reconstruction was intended for the S_0 mode, but the frequencies realized in the experiment proved to be outside the range for which the approximate solution used for the shape gradient is valid. Diligent et al. [120] and Grahn [88] describe models that go beyond the plane stress to account for normal variations in stress across the plate thickness.

Lamb wave propagation in plates is accurately described by the Helmholtz equation, where the field quantity is the displacement potential. While the experimental data were of very high quality, the boundary conditions associated with the displacement potential formulation of Lamb waves are incompatible with the shape gradient derived for the Helmholtz equation, except for limiting values of ka . Unfortunately, the experiment did not overlap these ka values, as it was specifically designed to span a range just above and below $ka = 1$.

The second major theme of this research applied optical flow to solving a new and unique problem in the field of synthetic aperture sonar. Areas of acoustic focusing and dilution attributable to refraction can sometimes resemble the natural bathymetry of the ocean floor. The difference is often visually indistinguishable, so it is desirable to

have a means of detecting these transient refractive effects without having to repeat the survey. The optical flow proved to be effective for this purpose.

Since any detection scheme would ideally be automated, connections were established between the tunable parameters of Horn and Schunck’s optical flow and the intrinsic properties of SAS imagery. Perhaps the most notable of these is the fact that the well-known statistical behavior of coherent speckle can be used to appropriately weight the smoothing term of the flow. The same principles could be applied to synthetic aperture radar imagery, which is known to be susceptible to similar effects under the right set of atmospheric conditions. Whether applied to sonar or radar, the potential exists for using optical flow as a tool to increase understanding of the dynamics of the ocean and the atmosphere.

7.1 Open Questions and Suggested Future Work

There are several clear avenues for continuing the investigation of adjoint active surfaces. The first is to further pursue experimental data that can be used to validate the approach and more firmly establish it as a viable alternative to existing techniques. Such experiments might be carefully constructed to prove a point, as was the ultrasonics experiment presented here. Or, data sets that work well with existing methods could be identified and compared with the active surfaces to better understand their relative strengths. Regardless, the best applications for adjoint active surfaces may not be the ones currently enjoying the most attention from the community.

An outstanding question is the precise nature of the relationship between PDE-based processing and classical array processing, including superresolution techniques. It is also not yet clear what role bandwidth plays in PDE-based reconstruction and what is the best way to use it. For traditional imaging methods bandwidth normally translates into resolution. For active surfaces, the utility of bandwidth appears to be in resolving ambiguities, that is, sorting out the globally optimal point in the cost

function from surrounding local extrema.

Optical flow has been experimentally demonstrated to show promise for detecting the presence of transient refractive effects using a single SAS image. The next logical step would be to fully automate the detection since the visual cues hinting at the presence of refractive anomalies can be subtle, often mimicking the appearance of natural effects. Such a capability would be useful for identifying portions of a survey area that may suffer from reduced image quality. A far more ambitious goal is to measure the character of the refracting disturbance, for example, the shape of its boundary and the sound speed profile within it. Such measurements could have a significant impact on the science of oceanography. This is also a point of intersection with adjoint active surface reconstruction, which could in principle be used to infer the properties of the refracting portion of the medium based on the changing backscatter from the sea floor.

REFERENCES

- [1] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar Systems and Signal Processing*. John Wiley and Sons, 1991.
- [2] W. G. Carara, R. S. Goodman, and R. Majewski, *Spotlight Synthetic Aperture Radar: Signal Processing Algorithms*. Norwood, MA: Artech House, Inc., 1995.
- [3] C. V. Jakowatz, D. E. Wahl, P. H. Eichel, D. Ghiglia, and P. Thompson, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*. Boston, MA: Kluwer Academic Publishers, 1996.
- [4] D. W. Hawkins, *Synthetic Aperture Imaging Algorithms: with application to wide bandwidth sonar*. PhD thesis, Department of Electrical and Electronic Engineering, University of Canterbury, Christchurch, New Zealand, 1996.
- [5] D. A. Cook, “Synthetic aperture sonar motion estimation and compensation,” Master’s thesis, Georgia Institute of Technology, 2007.
- [6] M. Pinto and A. Bellettini, “Shallow water synthetic aperture sonar: an enabling technology for NATO MCM forces,” Tech. Rep. NURC-PR-2007-010, NATO Undersea Research Centre, 2007.
- [7] K. Aki and P. G. Richards, *Quantitative Seismology, Theory and Methods*. Freeman, 1980.
- [8] A. Fichtner, *Full Seismic Waveform Modelling and Inversion*. Springer-Verlag, 2010.
- [9] D. Boiero and V. Socco, “Retrieving lateral variations from surface wave dispersion curves,” *Geophysical Prospecting*, vol. 58, pp. 221–233, 2010.
- [10] J. Virieux and S. Operto, “An overview of full-waveform inversion in exploration geophysics,” *Geophysics*, vol. 74, no. 6, pp. 127–152, 2009.
- [11] S. Bignardi, F. Fedele, G. Rix, A. Yezzi, and G. Santarato, “Geometric seismic wave inversion by the boundary element method,” *Bulletin of the Seismological Society of America*, vol. 102, pp. 802–811, April 2012.
- [12] F. Fedele, J. P. Laible, and M. J. Eppstein, “Coupled complex adjoint sensitivities for frequency-domain fluorescence tomography: theory and vectorized implementation,” *Journal of Computational Physics*, vol. 187, no. 2, pp. 597–619, 2003.

- [13] M. J. Eppstein, F. Fedele, J. P. Laible, C. Zhang, A. Godavarty, and E. M. Sevick-Muraca, “A comparison of exact and approximate adjoint sensitivities in fluorescence tomography,” *IEEE Transactions on Medical Imaging*, vol. 22, no. 10, pp. 1215–1223, 2003.
- [14] F. Fedele, M. J. Eppstein, J. P. Laible, A. Godavarty, and E. M. Sevick-Muraca, “Fluorescence photon migration by the boundary element method,” *Journal of Computational Physics*, vol. 210, pp. 109–132, 2005.
- [15] M. A. Richards, J. A. Scheer, and W. A. Holm, eds., *Principles of Modern Radar: Advanced Techniques*. SciTech Publishing, Inc., 2012.
- [16] N. Bleistein, J. K. Cohen, and J. W. S. Jr., *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*. Springer-Verlag, 2001.
- [17] M. A. Richards, “A beginner’s guide to interferometric SAR concepts and signal processing,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 22, September 2007.
- [18] R. Haberman, *Applied Partial Differential Equations with Fourier Series and Boundary Value Problems*. Pearson Prentice Hall, 2004.
- [19] M. G. Amin, “Radar, signal, and image processing techniques for through the wall imaging,” in *Proceedings of the SPIE*, vol. 5819, pp. 33–45, 2005.
- [20] M. G. Amin and F. Ahmad, “Wideband synthetic aperture beamforming for through-the-wall imaging,” *IEEE Signal Processing Magazine*, vol. 25, no. 4, pp. 110–113, 2008.
- [21] M. Cheney, “Mathematics and synthetic aperture radar imaging,” *SIAM News*, vol. 40, October 2007.
- [22] P. C. Chang, *Physics-Based Inverse Processing and Multi-path Exploitation for Through-Wall Radar Imaging*. PhD thesis, Ohio State University, 2011.
- [23] K. Chetty, G. E. Smith, and K. Woodbridge, “Through-the-wall sensing of personnel using passive bistatic wifi radar at standoff distances,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 1218–1226, April 2012.
- [24] J. Högbom, “Aperture synthesis with a non-regular distribution of interferometer baselines,” *Astronomy and Astrophysics Supplement*, vol. 15, 1974.
- [25] A. Blomberg and M. Hayes, “Multipath reduction for bathymetry using adaptive beamforming,” in *Proceedings of IEEE OCEANS 2010*, 2010.
- [26] J. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Prentice-Hall, 1993.
- [27] *IEEE Transactions on Geoscience and Remote Sensing, Special Issue on TerraSAR-X*, February 2010.

- [28] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping*. Wiley-Interscience, 1998.
- [29] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr., *Interferometry and Synthesis in Radio Astronomy*. Wiley, 2013.
- [30] J. R. Fineup, “Synthetic-aperture radar autofocus by maximizing sharpness,” *Optics Letters*, vol. 25, pp. 221–223, February 2000.
- [31] J. R. Fineup and J. J. Miller, “Aberration correction by maximizing generalized sharpness metrics,” *Journal of the Optical Society of America A*, vol. 20, pp. 609–620, April 2003.
- [32] T. M. Marston and D. S. Plotnick, “Semi-parametric statistical stripmap synthetic aperture autofocusing,” *IEEE Transactions on Geoscience and Remote Sensing* (Submitted), vol. 53, pp. 2086–2095, April 2015.
- [33] L. Gomez, M. E. Buemi, J. Jacobo-Berlles, and M. Mejail, “Introducing a new image quality index for evaluating despckling filtering operation in SAR images,” in *Proceedings of the 10th European Conference on Synthetic Aperture Radar*, 2014.
- [34] T. M. Benson, D. P. Campbell, and D. A. Cook, “Gigapixel spotlight synthetic aperture radar backprojection using clusters of GPUs and CUDA,” in *Proceedings of the 2012 IEEE Radar Conference*, pp. 853–858, 2012.
- [35] J. W. Goodman, *Statistical Optics*. John Wiley & Sons, 1985.
- [36] Y. Yu and S. T. Acton, “Speckle reducing anisotropic diffusion,” *IEEE Transactions on Image Processing*, vol. 11, pp. 1260–1270, November 2002.
- [37] D. A. Cook, “Synthetic aperture sonar speckle reduction using partial differential equation techniques,” in *Proceedings of the International Conference on Synthetic Aperture Sonar and Synthetic Aperture Radar*, Institute of Acoustics, 2006.
- [38] O. Germain and P. Réfrégier, “Edge location in sar images: Performance of the likelihood ratio filter and accuracy improvement with an active contour approach,” *IEEE Transactions on Image Processing*, vol. 10, pp. 72–78, January 2001.
- [39] J. C. Isaacs and R. Goroshin, “Tracking cables in sonar and optical imagery,” in *39th IEEE Applied Imagery Pattern Recognition Workshop*, 2010.
- [40] A. P. Lyons, R. E. Hansen, T. O. Saebø, and H. J. Callow, “Refractive effects of internal waves on synthetic aperture sonar images,” in *Proceedings of the 1st Underwater Acoustics conference, Corfu, Greece*, 2014.

- [41] F. M. Dickey, A. W. Doerry, and L. A. Romero, “Degrading effects of the lower atmosphere on long-range airborne synthetic aperture radar imaging,” *IET Radar, Sonar, and Navigation*, vol. 1, no. 5, pp. 329–339, 2007.
- [42] R. E. Hansen, A. P. Lyons, T. Sæbø, H. J. Callow, and D. A. Cook, “The effect of internal wave-related features on synthetic aperture sonar,” *IEEE Journal of Oceanic Engineering (In Press)*, 2014.
- [43] S. R. Doctor, T. E. Hall, and L. D. Reid, “SAFT – the evolution of a signal processing technique for ultrasonic testing,” *NDT International*, vol. 19, pp. 163–167, June 1986.
- [44] V. Schmitz, S. Chakhlov, and W. Müller, “Experiences with synthetic aperture focusing technique in the field,” *Ultrasonics*, vol. 38, pp. 731–738, 2000.
- [45] J. E. Michaels, S. J. Lee, J. S. Hall, and T. E. Michaels, “Multi-mode and multi-frequency guided wave imaging via chirp excitations,” in *Proceedings of SPIE, Health Monitoring of Structural and Biological Systems* (T. Kundu, ed.), no. 7984, 2010.
- [46] J. E. Michaels, J. S. Hall, G. Hickman, and J. Krolik, “Sparse array imaging of change-detected ultrasonic signals by minimum variance processing,” in *Review of Progress in Quantitative Nondestructive Evaluation* (D. O. Thompson and D. E. Chimenti, eds.), vol. 28A, pp. 642–649, American Institute of Physics, 2009.
- [47] J. L. Rose, *Ultrasonic Waves in Solid Media*. Cambridge University Press, 2004.
- [48] C.-C. Mow and Y.-H. Pao, “Diffraction of elastic waves and dynamic stress concentrations,” tech. rep., Rand Corporation, 1971.
- [49] J. F. P. J. Abascal, M. Lambert, D. Lesselier, and O. Dorn, “3-D eddy-current imaging of metal tubes by gradient-based, controlled evolution of level sets,” *IEEE Transactions on Magnetics*, vol. 44, no. 12, 2008.
- [50] O. Dorn, “A transport-backtransport method for optical tomography,” *Inverse Problems*, October 1998.
- [51] O. Dorn, E. Miller, and C. Rappaport, “A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets,” *Inverse Problems*, vol. 16, no. 5, 2000.
- [52] A. D. Zacharopoulos, S. R. Arridge, O. Dorn, V. Kolehmainen, and J. Sikora, “Three-dimensional reconstruction of shape and piecewise constant region values for optical tomography using spherical harmonic parametrization and a boundary element method,” *Inverse Problems*, vol. 22, pp. 1509–1532, 2012.

- [53] L. J. Cutrona, “Comparison of sonar system performance achievable using synthetic-aperture techniques with the performance achievable by more conventional means,” *Journal of the Acoustical Society of America*, vol. 58, no. 2, pp. 336–348, 1975.
- [54] M. P. Bruce, “A processing requirement and resolution capability comparison of side-scan and synthetic-aperture sonars,” *IEEE Journal of Oceanic Engineering*, vol. 17, no. 1, pp. 106–117, 1992.
- [55] R. W. Sheriff, “Synthetic aperture beamforming with automatic phase compensation for high frequency sonars,” in *Proceedings of the 1992 Symposium on Autonomous Underwater Vehicle Technology*, pp. 236–245, IEEE, 1992.
- [56] D. C. Munson, Jr., J. D. O’Brien, and W. K. Jenkins, “A tomographic formulation of spotlight-mode synthetic aperture radar,” *Proceedings of the IEEE*, vol. 71, pp. 917–925, August 1983.
- [57] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Society of Industrial and Applied Mathematics, 2001.
- [58] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances, Second Edition*. SPIE Press, 2009.
- [59] R. H. Stolt and A. B. Weglein, *Seismic Imaging and Inversion: Volume 1: Application of Linear Inverse Theory*. Cambridge University Press, 2012.
- [60] R. Stolt, “Migration by Fourier transform,” *Geophysics*, vol. 43, pp. 23–48, 1978.
- [61] C. Cafforio, C. Prati, and F. Rocca, “SAR data focussing using seismic migration,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 27, pp. 199–207, 1991.
- [62] A. J. Devaney, *Mathematical Foundations of Imaging, Tomography and Wavefield Inversion*. Cambridge University Press, 2012.
- [63] J. Dominguez, *Boundary elements in dynamics*. Computational Mechanics, 1993.
- [64] O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method: Volume 1 Basic Formulations and Linear Problems*. McGraw-Hill, 1989.
- [65] G. Fairweather and A. Karageorghis, “The method of fundamental solutions for elliptic boundary value problems,” *Advances in Computational Mathematics*, vol. 9, no. 1, pp. 69–95, 1998.
- [66] R. M. Errico, “What is an adjoint model?,” *Bulletin of the American Meteorological Society*, 1997.

- [67] P. Hursky, M. Porter, B. D. Cornuelle, W. S. Hodgkiss, and W. A. Kuperman, “Adjoint modeling for acoustic inversion,” *Journal of the Acoustical Society of America*, vol. 115, February 2004.
- [68] J. Tromp, C. Tape, and Q. Liu, “Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels,” *Geophysical Journal International*, vol. 160, pp. 195–216, 2005.
- [69] I. M. Ross, *A Primer on Pontryagin’s Principle in Optimal Control*. Collegiate Publishers, 2009.
- [70] J. X. Lopez, “Inverse synthetic aperture radar imaging theory and applications,” Master’s thesis, University of Texas Pan-American, 2011.
- [71] G. I. Marchuk, *Adjoint Equations and Analysis of Complex Systems*. Kluwer Academic Publishers, 1995.
- [72] D. J. Estep, “A short course on duality, adjoint operators, Green’s functions, and a posteriori error analysis,” 2011.
- [73] A. Sommerfeld, *Partial Differential Equations in Physics*. New York, New York: Academic Press, 1949.
- [74] N. Bleistein, *Mathematical Methods for Wave Phenomena*. Academic Press, 1984.
- [75] A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications*. Acoustical Society of America, 1989.
- [76] G. I. Marchuk, A. I. Agoshkov, and V. P. Shutyaev, *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems*. CRC Press, 1996.
- [77] N. Sun, *Inverse Problems in Groundwater Modeling*. Kluwer Academic Publishers, 1994.
- [78] S. J. Norton, “Iterative inverse scattering algorithms: Methods of computing,” *Journal of the Acoustical Society of America*, vol. 106, November 1999.
- [79] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*. Princeton University Press, 1987.
- [80] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*. Elsevier, 2004.
- [81] J. D. Jackson, *Classical Electrodynamics*. John Wiley and Sons, 1975.
- [82] D. Nain, A. Yezzi, and G. Turk, “Vessel segmentation using a shape driven flow,” in *Lecture Notes in Computer Science: Medical Image Computing and Computer-Assisted Intervention*, 2004.

- [83] L. He, S. Kindermann, and M. Sini, “Reconstruction of shapes and impedance functions using few far-field measurements,” *Journal of Computational Physics*, vol. 228, no. 3, pp. 717–730, 2009.
- [84] A. Litman, D. Lesselier, and F. Santosa, “Reconstruction of a two-dimensional binary obstacle by controlled evolution of a level-set,” *Inverse Problems*, vol. 14, no. 3, 1999.
- [85] M. Eskandari and R. Safian, “Inverse scattering method based on contour deformations using a fast marching method,” *Inverse Problems*, vol. 26, no. 9, 2010.
- [86] I. A. Viktorov, *Rayleigh and Lamb Waves: Physical Theory and Applications*. Plenum Press, 1967.
- [87] J. D. Achenbach, *Wave Propagation in Elastic Solids*. North-Holland, 1975.
- [88] T. Grahn, “Lamb wave scattering from a circular partly through-thickness hole in a plate,” *Wave Motion*, vol. 37, pp. 63–80, 2015.
- [89] Y.-H. Pao, “Dynamical stress concentration in an elastic plate,” *Journal of Applied Mechanics*, pp. 299–305, June 1962.
- [90] D. R. Christman, W. M. Isbell, S. G. Babcock, A. R. McMillan, and S. J. Green, “Measurements of Dynamic Properties of Materials Volume III: 6061-T6 Aluminum,” tech. rep., General Motors Corporation, 1971.
- [91] C. S. Clay and H. Medwin, *Acoustical Oceanography*. John Wiley and Sons, Inc., 1977.
- [92] M. A. Richards, J. A. Scheer, and W. A. Holm, eds., *Principles of Modern Radar: Basic Principles*. SciTech Publishing, Inc., 2010.
- [93] K. F. Graff, *Wave Motion in Elastic Solids*. Dover Publications, 1991.
- [94] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*. Prentice-Hall, 1983.
- [95] J. H. Nieuwenhuis, J. J. Neumann, D. W. Greve, and I. J. Oppenheim, “Simulation and testing of transducers for lamb wave generation,” in *2005 IMAC-XXIII: Conference & Exposition on Structural Dynamics*, 2005.
- [96] D. A. Cook, M. Mueller, F. Fedele, and A. J. Yezzi, “Adjoint active surfaces for localization and imaging,” *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 316–331, 2015.
- [97] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Inc., 1983.
- [98] “NIST Digital Library of Mathematical Functions, NIST-DLMF.”

- [99] M. Avriel, *Nonlinear Programming: Analysis and Methods*. Prentice-Hall, 1976.
- [100] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*. Springer, 2008.
- [101] J. C. Lagaria, J. A. Reeds, M. H. Wright, and P. E. Wright, “Convergence properties of the Nelder-Mead simplex method in low dimensions,” *SIAM Journal of Optimization*, vol. 9, no. 1, pp. 112–147, 1998.
- [102] K. R. Helfrich, “Internal solitary wave breaking and run-up on a uniform slope,” *Journal of Fluid Mechanics*, pp. 133–154, 1992.
- [103] J. B. Hersey, ed., *Deep-Sea Photography*. Johns Hopkins Press, 1967.
- [104] B. C. Heezen and C. D. Hollister, *The Face of the Deep*. Oxford University Press, 1971.
- [105] P. C. Wille, *Sound Images of the Ocean in Research and Monitoring*. Springer-Verlag, 2005.
- [106] J. R. Apel, L. A. Ostrovsky, Y. A. Stepanyants, and J. F. Lynch, “Internal solitons in the ocean,” Tech. Rep. WHOI-2006-04, Woods Hole Oceanographic Institution, 2006.
- [107] S. K. Venayagamoorthy and O. B. Fringer, “On the formation and propagation of nonlinear internal boluses across a shelf break,” *Journal of Fluid Mechanics*, pp. 137–159, 2007.
- [108] D. Bourgault, M. D. Blokhina, R. Mirshak, and D. E. Kelley, “Evolution of a shoaling internal solitary wave train,” *Geophysical Research Letters*, vol. 34, 2007.
- [109] B. R. Munson, D. F. Young, and T. H. Okiishi, *Fundamentals of Fluid Mechanics*. John Wiley & Sons, Inc., 2 ed., 1994.
- [110] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer-Verlag, 2002.
- [111] B. K. P. Horn and B. G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.
- [112] H. Goldstein, *Classical Mechanics*. Addison-Wesley Publishing Company, Inc., 1953.
- [113] L. C. Evans, *Partial Differential Equations*. American Mathematical Society, 2002.
- [114] T. O. Sæbø, *Seafloor Depth Estimation by Means of Interferometric Synthetic Aperture Sonar*. PhD thesis, University of Tromsø, Norway, 2010.

- [115] H. A. Zebker and J. Villasenor, “Decorrelation in interferometric radar echoes,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 950–959, September 1992.
- [116] D. C. Munson, Jr. and J. L. C. Sanz, “Image reconstruction from frequency-offset Fourier data,” *Proceedings of the IEEE*, vol. 72, pp. 661–669, June 1984.
- [117] J. W. Goodman, *Speckle Phenomena in Optics*. Roberts and Company Publishers, 2010.
- [118] W. F. Walker and G. E. Trahey, “The application of k-space in pulse echo ultrasound,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 45, no. 3, 1998.
- [119] J. Groen, R. E. Hansen, H. J. Callow, J. C. Sabel, and T. O. Sæbø, “Shadow enhancement in synthetic aperture sonar using fixed focusing,” *IEEE Journal of Oceanic Engineering*, vol. 34, pp. 269–284, July 2009.
- [120] O. Diligent, T. Grahn, A. Boström, P. Cawley, and M. J. S. Lowe, “The low-frequency reflection and scattering of the s_0 Lamb mode from a circular through-thickness hole in a plate: Finite element, analytical and experimental studies,” *Journal of the Acoustical Society of America*, vol. 112, pp. 2589–2601, December 2002.