Travel-Time Tomography for Stress Reconstruction in Granular Soil Media

by

Abbas Mohammad Vali Samani

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 1997

©Abbas Mohammad Vali Samani 1997

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-22221-7

Canada

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

Computerized tomography (CT) has long been used successfully in different areas of science and engineering. X-ray tomography and MRI in diagnostic medicine, and geographical imaging, usually used for characterizing the subsurface of earth, are well-known examples. For experimental determination of stress fields, transmission photoelasticity has been well established and successfully applied in solid bodies. However, for the determination of stress fields in particulate media, especially in granular soil medium, very little research has been conducted. Here, based on CT principles, a non-destructive technique is presented for the determination of stress fields in granular soil medium. This technique is based on crosshole seismic probing, which is usually used by geophysicists for geophysical imaging. In this technique, a set of seismic energy sources and a set of sensors are distributed on different boundaries of the soil region of interest. The measured propagation delay between each source-receiver pair, usually called the "travel-time", is used to infer the stress field.

The wave velocity field in granular soil medium is essentially anisotropic and heterogeneous. In conjunction with the borehole technique, two different mathematical models have been presented to approach the problem. One is based on the states of stress at each point in the domain as the principal unknowns to be determined, and the other is based on defining the stress field as a mathematical expression with a certain number of unknown parameters. Both models lead to a constrained nonlinear least squares problem. However, the first turned out to be technically unmanageable due to the highl nonlinearity of the governing equations and the large number of unknowns involved, while the latter proved to be not only manageable but also an efficient approach. To demonstrate the merits of this approach for the determination of stress field of granular soil, a number of test cases

corresponding to different boundary conditions are presented where the stress fields are reconstructed from travel-times which are produced numerically.

Acknowledgements

I would like to thank my supervisor, Prof. G.M.L. Gladwell, for his encouragement and support, and for being source of inspiration throughout my Ph.D. program.

I owe a debt of thanks to Prof. G.M.L. Gladwell, and Prof. J.C. Santamarina whose courses "Inverse Problems" and "Wave-Material Interaction and Applications" formed the foundation for much of my postgraduate work.

I would like to thank C. Campbell for his valuable assistance in debugging my codes. Also, I would like to thank Miss Hanene Ben-Abdallah for typing the majority of this thesis.

This work would have not been possible without the financial support of the ministry of Culture and Higher Education of the Islamic Republic of Iran. This support is highly appreciated and is hereby acknowledged. Last, but not least, I would like to thank my wife Mitra and my little beautiful daughters Atefeh and Negar for giving me the love and understanding that made this Ph.D. possible.

Finally, and most of all, I am thankful to the Almighty God for all blessing bestowed upon me throughout my Ph.D. program.

Contents

1 Introduction 1

1.1 General Background to the Problem 1

1.2 General Background 2

1.3 Research Objective and Thesis Preview 6

2 Forward Modeling in CT 10

2.1 Introduction 10

2.2 The Forward Problem 11

2.2.1 Wave Propagation in Particulate Media 12

2.2.2 Anisotropy in Granular Soil 12

2.2.3 Wave Velocity in Cross-Anisotropic Media 15

2.2.4 Velocity-Stress Relations in Particulate Media: Micromechanics Based Relations 17

2.2.5 Wave Approach in Forward Modeling 20

2.2.6 Reduced Wave Equation Approach of Forward Modeling . . 21

2.2.7 Ray Based Approximation of Forward Modeling 23

2.3 Closure 30

3 Inverse Problem in CT **32**

3.1 Introduction 32

3.2 Inverse Problem in the Wave Approach 32

3.3 Transform Based Inversion Methods 34

3.4 Ray Based Travel-Time Tomography 35

3.4.1 Reconstruction Techniques in Straight-Ray Tomography . . 36

3.4.2 Iterative Schemes in Curved-Ray Tomography 41

3.4.3 Ray Based Travel-Time Tomography in Inhomogeneous Anisotropic Media 48

3.5 The Nonlinear Least-Squares Problem 49

3.6 Large Scale Sparse Nonlinear Least-Squares Problems 53

3.7 Constrained Optimization 54

3.7.1 Penalty-Function Methods 56

3.7.2 Barrier Function Method 57

3.7.3 Generalized Reduced-Gradient Methods 57

3.7.4 Sequential Quadratic Programming Methods 58

3.8 Constrained Nonlinear Least-Squares Problem 60

3.9 Sparse Least-Squares Problem 63

3.9.1 The Method of Normal Equations 63

3.9.2 QR Factorization 64
3.9.3 LSQR Algorithm 65
3.9.4 Solving The Augmented System 67
3.10 Closure 67

4 Mathematical Modeling **70**

4.1 Introduction 70
4.2 Ray Based Travel-Time Equation for Stress Induced Anisotropic-Inhomogeneous Media 72
4.3 Ray Tracing in Stress Induced Inhomogeneous Anisotropic Media . 74
4.3.1 A Forward-Backward Type Ray Tracing Method 74
4.3.2 Fermat's Principle-Based Ray Tracing Method 78
4.4 The Inverse Problem 79
4.4.1 Nodal Stress Approach 79
4.4.2 Stress Potential Function Approach 84
4.5 A Method for the Velocity-Stress Power Relationship Parameters Determination 88
4.6 The Stress Potential Function Approach in 3-D 89
4.6.1 Travel-Time Equation in 3-D 89
4.6.2 Ray Tracing in Stress Induced Inhomogeneous Anisotropic Media 91
4.6.3 The Inverse Problem 92

5 Numerical Solution of the Model **94**

5.1 Introduction 94

5.2 Numerical Formulation 95

5.2.1 Ray Tracing Using the Simplex Method 96

5.3 Numerical Formulation of the Inverse Problem 100

5.3.1 Nodal Stress Approach: Numerical Formulation 100

5.3.2 Stress Function Approach: Numerical Formulation 111

5.4 Scaling 114

6 Numerical Examples **118**

6.1 Introduction 118

6.2 Ray Tracing in a Soil Medium Under Footing 119

6.3 Stress Reconstruction in Footing Problems 123

6.3.1 Footing Problem Stress Reconstruction: Nodal Stress Approach 123

6.3.2 Footing Problem Stress Reconstruction: Stress Functions Approach 125

6.3.3 Tunnel Problem Stress Reconstruction: Stress Functions Approach 131

6.3.4 Retaining Wall Problem Stress Reconstruction: Stress Functions Approach 137

6.4 Closure 142

7 Closure **143**

Bibliography 145

List of Figures

1.1 Subsurface region bounded by boreholes 4

2.1 Relationship between the wave surface and phase velocity surface . 15

3.1 SLIM-Straight line inversion with modified data 43

3.2 CRAI-Curved ray algebraic inversion 45

4.1 Elliptical anisotropy of slowness 73

4.2 A typical grid used in ray tracing 75

4.3 Segment AB of a ray path in a typical cell 76

4.4 Stress Components 81

4.5 Stress Components 92

5.1 The basic moves of the simplex method 97

5.2 Flow diagram of the ray tracing procedure 99

5.3 A rectangular cell with nodal points in a 2-D medium 101

5.4 A rectangular cell with linear variations under: (a) σ_x, (b) σ_y, (c) τ 103

5.5 Nodal parameters used in forming the regularization operator . . . 105

5.6 Flow Chart of the numerical solution in the nodal stress approach . 112

5.7 Flow Chart of the numerical solution in the stress potential function approach 115

6.1 Soil region under footing 120

6.2 Distribution of σ_x, σ_y, and τ of the soil medium under footing . . . 121

6.3 Ray paths between sources and receivers in a soil region under footing122

6.4 Soil region modeled as a half-space subjected to a uniform load . . 124

6.5 Distribution of σ_x, σ_y, and τ in the rectangular subregion 126

6.6 Distribution of σ_x, σ_y, and τ in the initial guess 127

6.7 Reconstructed distribution of σ_x, σ_y, and τ in the footing problem using the nodal stress approach 128

6.8 Footing problem reconstruction results 132

6.9 Soil region with a tunnel 133

6.10 Tunnel problem reconstruction results 136

6.11 Soil region supported by a retaining wall 137

6.12 Retaining wall problem reconstruction results 141

To:

My wife. who - more than I - endured the hardship of this work.

and my parents for their encouragement.

Chapter 1

Introduction

1.1 General Background to the Problem

Stress analysis of soil is one of the most challenging topics in geotechnical engineering. Due to the particulate nature, the real mechanical behavior of soil is very complex compared to that of continuous solid materials.

The Finite Element Method has long been used for stress analysis of a soil medium. In classical Finite Element models, soil is modeled as a continuum; to compensate for its particulate nature, the mechanical behavior is usually considered elasto-plastic with strain hardening [Che85]. The elasto-plasticity and strain hardening characteristics have a limited validity and are different for different types of soil; the results have a high level of uncertainty. This dissertation embodies the results of a research program that was motivated by a number of recent papers such as Santamarina et al [SGM+93], and Santamarina and Potts [SP94] which provided experimental evidence that field measurements of travel-time could be used to infer the state of stress in particulate media. That is, the objective of the present work

is to provide a new technique for the determination of stress fields in granular soil without incorporating any stress-strain constitutive law. This technique resembles Photoelasticity which is an optical experimental technique designed to determine stress fields in solid bodies with complicated geometry. In photoelasticity, light is passed through a loaded transparent model of the body, which is initially isotropic and homogeneous. The stress field caused by the loading causes the transparent medium to become optically anisotropic. The optical anisotropy causes alteration of the of the light characteristics such as changes in phase or polarization. These are measured and related to the stress field by means of a mathematical framework. In this technique, seismic waves are passed through a loaded region of soil that is initially assumed to be isotropic and homogeneous. The stress here similarly causes a non-homogeneous and anisotropic velocity field of the seismic waves. Alteration of the original wave characteristics such as propagation delay between pairs of points located on the boundary are measured and used in a mathematical model to predict the stress field.

This technique provides a proper tool for determining the stress field using data obtained from measurements *in situ*. The proposed technique can be applied to verify classical models of stress analysis or to diagnose whether an existing soil region such as an earth dam is close to failure. To apply this technique for design purposes in similar fashion to photoelasticity requires further research in model analysis.

1.2 General Background

Tomography is derived from the Greek word $\tau o \mu o \sigma$ (*tomos*) that is, slice; it also refers to the cross-sectional imaging of an object from either transmission or reflec-

tion data collected by illuminating the object from many different directions. The object is illuminated by means of a certain kind of energy that excites the object and is transmitted in the form of waves.

During transmission, waves do not alter the characteristics of the object significantly; therefore tomography is regarded as a nondestructive method of imaging. Since measurements of data are made on the boundary of the object, tomography provides a relatively convenient way for collecting the data required for imaging purposes.

The application of tomographic imaging in diagnostic medicine is well known, and the first medical application used x-rays for forming images of tissues based on their x-ray attenuation coefficients.

Geotomography is a more recent application of computerized tomography as a nondestructive method in geophysical exploration. In contrast to medical tomography where the object can be interrogated from all sides, in geotomography a region underground can be viewed from only a few sides. Surface to surface data collection is one of the most convenient strategies in geotomography, but it usually leads to poor images. Borehole to borehole methods are more appropriate. Additional surface or mine gallery sources and receivers where feasible gives better quality images. This work is concerned with the borehole to borehole case, usually called *crosshole tomography*, however the presented methodology is fundamentally general, and additional surface or mine gallery measurements can easily be incorporated. A subsurface region bounded by two boreholes including transmitters and receivers is shown in Figure 1.1.

In geophysical applications, seismic energy, usually produced by explosion, is used to illuminate the region. The energy propagation in the region is governed by

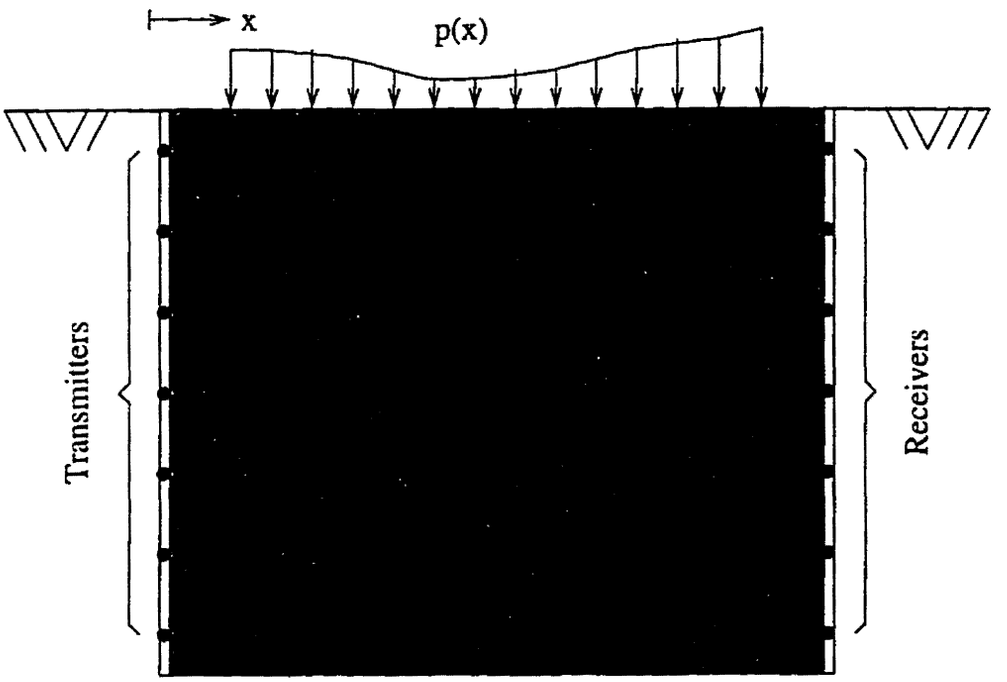


Figure 1.1: Subsurface region bounded by boreholes

the acoustic or the elastic wave equation. If the wavelength is much smaller than the size of inhomogeneities, the wave equation can be simplified to the well-known eikonal equation. If the conditions are not satisfied, diffraction must be considered. For instance, if we have an anomaly, particularly a low velocity anomaly, with a size comparable to the wavelength, diffraction should certainly be considered.

As waves are transmitted through the region, based on the distribution of certain parameters in the region, they are reflected, refracted and/or diffracted. The collected data contain information about these phenomena and therefore about the distribution of these parameters in the region. In the present work, the main concern is evaluating the stress distribution in the region of soil. We assume that there is no interface in the region which may cause reflection; the stress is supposed to change continuously, therefore there will be no significant diffraction. Consequently, in this work, refraction will be the main factor that is emphasized.

To obtain a rough idea about difficulties involved in geophysical tomography, we can consider the following equation which relates measured travel-times (t_i) between two points on the boundary of the region and the spatial slowness distribution of wave propagation ($S(r)$):

$$\int_{R_i} S(r)dl_i = t_i \qquad i = 1, 2, \cdots, k \tag{1.1}$$

where R_i represents the ray path between two points along which seismic energy propagates, and k is the number of rays taken into consideration.

In equation (1.1), R_i generally is unknown and depends on the slowness distribution. Therefore, the problem is nonlinear and computationally involved. To overcome this problem, iterative methods are applied; however, they are not always guaranteed to converge. The other serious difficulty is the lack of uniqueness of the solution for the equations. This arises due to the limited directional coverage

when sources and receivers are located in vertical boreholes. Noise is an inevitable problem, which contaminates the received signals and consequently leads to an inconsistent system of equations that has no exact solution. Although contaminated data can be processed for imaging purposes, it will still affect the quality of the image significantly, even though the magnitude of noise is usually small in geophysical applications.

Anisotropy is one of the most serious problems which geophysicists often confront. In geophysical exploration applications, geophysicists are usually concerned about inherent anisotropy caused by crystal orientation, particle orientation, crack orientation or lamination. In this study, we are concerned with stress induced anisotropy caused by contact orientation of particles as a result of loading. Wave propagation in anisotropic media presents unique phenomena such as birefringence or S-wave splitting, and quasi-propagation, i.e., particle motions in the three quasi-body-waves will not be parallel to the corresponding displacements of P, S_v and S_h waves in isotropic media [Hel94, Aul73]. To account for the anisotropy, more unknowns must be introduced, consequently the already ill-conditioned problem becomes worse.

1.3 Research Objective and Thesis Preview

The objective of this work is to develop a method for stress determination in granular soils, based on tomographic principles. It is assumed that data is available which represents the behavior of acoustic waves in the soil. This data, in general, can be travel-times, wave amplitudes and/or full patterns of the output wave. In this work, we will focus on travel-time of P-waves as a representation of wave behavior. The topic of the reconstruction of stress distribution in granular soil using

measured travel-times is an inverse problem. Due to the novelty of this topic, to our knowledge, there is no mathematical model available that relates the measured travel-times to the stress distribution in granular soil media. Therefore, not only has the inverse problem not been investigated before, but also there has been only few attempts to tackle the forward problem, In an attempt to accomplish our objective, we carried out our research step by step as follows:

1. *Mathematical modeling.* Considering the heterogeneity and the stress induced anisotropy of the problem, and applying the existing relations between the velocity of wave propagation and stress in particulate medium, we have established an integral equation that relates the measured travel-times to the stress field. Unlike equation (1.1), our equation is nonlinear, not only due to its dependence on the ray paths, but also because of the nonlinear integrand. This unfortunately adds a significant level of complexity.

2. *Forward problem solution.* We have carefully studied ray tracing in isotropic and inherently anisotropic media, and concluded that a perturbation method that has been widely used in heterogeneous-isotropic media can suit our model. Therefore, we adopt this method and present an efficient ray tracing technique for stress induced heterogeneous-anisotropic soil media.

3. *Inverse problem solution.* First, we present an inversion approach that is based on using stresses as the principal unknowns. Due to the large number of unknowns, and the limited information which is provided from the travel-times because of their limited directional coverage, the inverse problem is extremely ill-conditioned and the stress field cannot be reconstructed properly, even by applying regularization. Therefore, we introduce the equilibrium equations which must govern the stress field. This dramatically decreases

the degree of instability of the equations; however, the problem remains ill-conditioned. To improve the stability of the problem further, we apply a regularization technique.

The system of equations derived from the regularization technique has to be dealt with as a nonlinear least-squares problem, and solved iteratively. On the other hand, considering the fact that, typically, the number of unknowns involved in this approach is very large, an efficient way to handle the iterative procedure of inverse problem is certainly required. In our model, fortunately, each governing equation of the system involves only a few unknowns, consequently its related Jacobian matrix is sparse. Therefore, an algorithm for solving sparse nonlinear least-squares problems was adopted and encoded to be applied in this problem. Unfortunately, due to the high nonlinearity of the travel-time equations and the large number of unknowns involved, test problems with synthetic data revealed that unless the initial guess is very close to the solution, the iterations often either diverge or converge to a local minimum.

As a second attempt to solve the inverse problem, we present another approach that is based on using stress functions with a limited number of unknowns to avoid some complexities we had in the first approach. The stress functions are selected in such a way that they automatically satisfy the equilibrium requirements, therefore the governing equations in this approach are only travel-time equations. These equations are stable, but still highly nonlinear. To avoid convergence to a local minimum in this approach, we impose a number of realistic constraints to this system of equations. Consequently the obtained system has to be dealt with as a constrained nonlinear least-squares problem which is one of the challenging problems in optimization.

This approach is encoded and applied to a number of test problems for the reconstruction of stress fields from synthetic travel-times.

The details of these approaches are contained in subsequent chapters. In chapters 2 and 3, we present a comprehensive literature survey on some relevant principles, methods and tools used in inverse problem applications which are to a certain degree similar to the problem of this study; Chapter 2 is devoted to issues related to forward modeling, while Chapter 3 is devoted to inverse problem issues.

In Chapter 4, the principles of our approach to solve the problem in 2-D cases are addressed, and a relation between travel-times and stress field is derived. Using this relation, we present two approaches to formulate the inverse problem. In this first approach, we represent the unknown stress field by considering the unknown stress components on a set of pre-specified nodes in the soil region of interest. In the second approach, we use the idea of stress potential functions to represent the stress field of the entire region parametrically. Compared to the first approach, the second has some attractive features such as involving only a small number of unknowns, and also involving a parametric stress field that automatically satisfies force equilibrium conditions. The last part of Chapter 4 essentially formulates the problem in 3-D cases. This formulation is basically an extension of the 2-D case formulation obtained from the second approach.

Chapter 5 addresses the numerical solution of both the forward and the inverse problems in details. In Chapter 6, we present the application of the model for stress reconstruction in different geotechnical problems. Finally, Chapter 7 is devoted to a summary and the conclusions of this work, along with a discussion of future possible work.

Chapter 2

Forward Modeling in CT

2.1 Introduction

The present study is probably one of the first attempts to apply tomography to stress distribution reconstruction in soil. Due to the novelty of this subject, to the author's knowledge, no computational model for stress distribution reconstruction has been presented in the literature so far. In this chapter, we will mainly review methods of forward modeling in computerized tomography, which are applied in geophysical explorations. These methods are developed to determine spatial distribution of wave propagation velocity in a region excited by seismic energy. These investigations are relevant to the present study because the foundations are basically the same.

The problem of determining the wave propagation velocity in a region can be approached in three ways. First is the wave approach which is based on the governing partial differential equation of wave propagation; fundamentally it is the most accurate approach and it appears to be widely applicable, however it requires a

cumbersome computational process that limits its application. The second, diffraction tomography, is based on a version of the wave equation that has been simplified to make the computational process of inversion easier. The simplifications are valid only in specific circumstances. The third set, ray based methods of tomography, are based on ray theory, which considers the transmission of energy along rays. This is valid only in even more specific circumstances, but is computationally the simplest.

Although the validity of diffraction tomography and ray based methods of tomography is limited, these methods have been widely used by researchers. This is due to their relative simplicity, applicability and accuracy in comparison to full wave approaches. Full wave approaches are theoretically ideal, but lead to severe computational difficulties.

Because the reconstruction model of this work is based on a ray approach, we will briefly discuss the foundations of wave and diffraction tomography approaches in forward modeling, and present a detailed discussion of ray based methods.

Tomography is an inverse problem. To solve any inverse problem, we must have a deep understanding of its corresponding forward problem. This Chapter is devoted to a review of different approaches and some other issues related to the forward modeling, while Chapter 3 is devoted to a review of techniques for solving the inverse problem.

2.2 The Forward Problem

In forward problems, the goal is to compute some behavior of a system corresponding to some parameter distribution. In geophysics, the primary forward problem is

the computation of travel-times for waves passing through a medium with known properties.

2.2.1 Wave Propagation in Particulate Media

The consideration of mechanical wave propagation in particulate media was first initiated by Newton in the 17^{th} century, when he modeled air as a discrete medium for propagation of sound [Alo95]. Wave propagation in particulate media is very complicated when the wavelength approaches the particle diameter. Brillouin, in his classical text [Bri46], has summarized the most significant developments related to this subject until the 1940's. One of the most important conclusions of Brillouin's work is that when the wavelength is much greater than the interparticle spacing, waves travel through particulate media as in a continuum. However, if the wavelength approaches the interparticle spacing, waves are affected by the particulate nature of the medium.

2.2.2 Anisotropy in Granular Soil

There are basically two kinds of soil anisotropy: inherent anisotropy which can be caused by a variety of ordering processes on a "microscopic" level, such as grain or crystal orientation, aligned cracks or fractures or periodic fine layering ("lamellation") [Hel94]; and stress induced anisotropy which is caused by interparticle contact orientation as a consequence of loading. Researchers in different fields are interested in different parameters related to anisotropy. For example, exploration geophysicists are interested in the causes that makes the medium deviate from isotropy, while most of civil engineers are interested in the elastic constants in anisotropic medium. In this research, we are interested in stress distribution in

a soil medium, therefore we are concerned with stress induced anisotropy and how it affects wave propagation

Stress Induced Anisotropy and Wave Propagation

When a soil region undergoes loading, equilibrium and compatibility requirements lead to an inhomogeneous-anisotropic stress distribution. The state of a particulate medium can be described at the micromechanical level by characterizing the distribution of contacts, interparticle forces and particle orientations. Hence, it is theoretically possible to relate the micromechanical characteristics of the particulate medium to the macro-characteristics of wave propagation such as velocity. Numerical and experimental micromechanics research of particulate media has led to important relations between the distribution of externally applied forces acting at the boundaries of a granular assembly, and the fabric anisotropy and anisotropic distribution of interparticle forces that evolve in response to boundary disturbances: the average number of contacts and the average normal contact forces increase in the direction of the higher normal stress until limiting values of force and fabric anisotropy are reached. This change arises as a consequence of particle arrangement and the distribution of interparticle contact orientations and contact forces in a particulate medium. Rothenburg [Rot80], and Rothenburg and Bathurst [RB89, RB92] have proposed a truncated Fourier series to describe the distribution of these parameters. Rothenburg and Bathurst [RB89, RB92] suggested the following approximation for the angular distribution of contact normals:

$$E(\theta) = \frac{1}{\pi}\{1 + a\cos 2(\theta - \theta_a)\} \tag{2.1}$$

where a is a parameter defining the magnitude of anisotropy in contact orientations and θ_a defines the direction of anisotropy. A similar expression has been suggested

for contact forces. Although it seems to be feasible to pursue this micromechanical approach to investigate wave propagation in particulate media with stress induced anisotropy, to the best of our knowledge, this approach has not been extended for this purpose.

Cross-Anisotropic Media

If a solid medium exhibits symmetry with respect to arbitrary rotations about one axis, then it is called *cross anisotropic*. In this case the matrix of elastic coefficients has only five dependent constants. Thus, with the z-axis as the axis of symmetry, Hooke's law can be expressed as follows:

$$\begin{Bmatrix} \sigma_x \\ \sigma_y \\ \tau_{yz} \\ \tau_{xz} \\ \tau_{xy} \end{Bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{bmatrix} \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \\ \gamma_{yz} \\ \gamma_{xz} \\ \gamma_{xy} \end{Bmatrix} \tag{2.2}$$

Inherently anisotropic soil is often modeled as cross-anisotropic medium especially in geophysical exploration applications [Č72, BC90, CP92, PC92]. In the case of stress induced anisotropy, depending on the loading conditions, soil may behave as a cross-anisotropic medium or a medium with one plane of symmetry. For example, when a soil medium undergoes a symmetric loading such as a circular footing, soil behaves as a cross-anisotropic medium, while under plane strain conditions, soil may be regarded as a medium with one plane of symmetry. In such cases it may be theoretically possible to develop relations between the elastic constants and the state of stress such as the one derived by Petrakis and Dorby [PD87]; such

relations, theoretically, can be used to derive the wave equation in the stress induced anisotropy case. However, the resulting wave equation would be extremely nonlinear, thus very hard to handle for further applications.

2.2.3 Wave Velocity in Cross-Anisotropic Media

A fundamental aspect of an anisotropic medium is the variation in velocity of a seismic wave with the direction of propagation. In anisotropic media, the velocity of plane wave propagation, called *phase velocity*, differs from that of wave propagation which is called ray velocity. As depicted in Figure 2.1, the wave surface is defined as the locus of points reached at a given time by an impulsive point source at time zero. Therefore a radial vector connecting the source to a point on this surface represents the ray velocity at that point. On the other hand, the phase velocity surface is defined as the locus of radial vectors, which are perpendicular to the plane wave fronts of the wave surface.

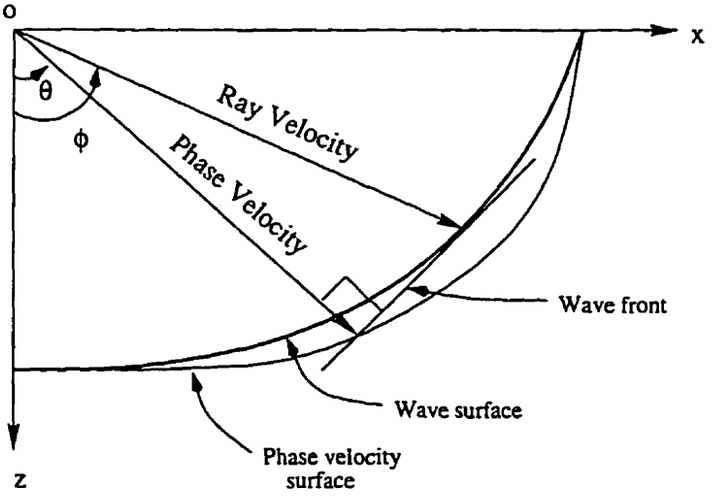


Figure 2.1: Relationship between the wave surface and phase velocity surface

Although no research has been conducted about wave propagation in media with one plane of symmetry, it is expected that the same phenomena exist.

Wave Equation Solution

The solution of wave equation for cross-anisotropic media [Whi83, ?] where the z-axis is the axis of symmetry leads to the following equations for phase velocities:

$$V_{p,\theta} = [(c_{11}\sin^2\theta + c_{44} + + c_{33}\cos^2\theta + \Delta)/(2\rho)]^{1/2} \quad (2.3)$$

$$V_{sh,\theta} = [c_{66}\sin^2\theta + c_{44}\cos^2\theta)/\rho]^{1/2} \quad (2.4)$$

$$V_{sv,\theta} = [(c_{11}\sin^2\theta + c_{33}\cos^2\theta + c_{44} - \Delta)/(2\rho)]^{1/2} \quad (2.5)$$

where

$$\Delta = [\{(c_{11} - c_{44})\sin^2\theta - (c_{33} - c_{44})\cos^2\theta\}^2 + 4(c_{13} + c_{44})^2\sin^2\theta\cos^2\theta]^{1/2}$$

and θ is the phase angle measured from the z-axis of symmetry. The P-wave ray velocity cannot be easily expressed in terms of stiffness coefficients in closed form and it can be determined through parametric representations [Byu84], however approximate models are available as follows:

4th Fourier Series Model. Byun et al [BCG89] proposed an approximate direct expression through a truncated Fourier type expansion of the form

$$V_P^{-2}(\theta) = a_0^2 + a_1^2\cos^2\theta - a_2^2\cos^4\theta \quad (2.6)$$

to describe the P-wave ray velocity surface for cross-anisotropic media. The three parameters a_0, a_1 and a_2 are functions of our stiffness coefficients c_{11}, c_{13}, c_{33} and c_{44}. Byun et al [BCG89] showed that, this three-term approximation to the ray velocity is an excellent fit to P-wave ray velocity even for highly anisotropic media.

Elliptical Model. In this case of cross-anisotropy, the ray velocity of SH-waves exhibits an elliptical dependence [Byu82, BC90]. In this case, $V_{\mathrm{sh},\theta}$ the SH-wave ray velocity as a function of ray angle θ can be expressed as

$$\frac{1}{V_{\mathrm{sh},\theta}^2} = \frac{\cos^2\theta}{V_v^2} + \frac{\sin^2\theta}{V_h^2} \tag{2.7}$$

which involves the vertical and horizontal velocities, respectively,

$$\begin{aligned} V_v &= \sqrt{\tfrac{c_{44}}{\rho}} \\ V_h &= \sqrt{\tfrac{c_{66}}{\rho}} \end{aligned}$$

Apparent P-wave anisotropy exhibited by some cross-anisotropic media is not strictly elliptical [Byu82], however, any convex part of the wave front can be locally fitted with an ellipse within the view angles involved in the specific field situation [Hel94]. Due to this fact, Michelena et al [MMH93] showed that P-wave ray velocity in cross-anisotropic media can be approximated properly by a double elliptical model that has elliptical shape around the axis of symmetry.

2.2.4 Velocity-Stress Relations in Particulate Media: Micromechanics Based Relations

Velocity-stress relations in particulate media can be derived using micromechanics principles. Applying Hertz contact theory for a cubic elastic sphere pack under isotropic confining stress σ_0 leads to the following equation for the elastic modulus [Whi83]:

$$E = \{\frac{3E_s^2\sigma_0}{8(1-\nu_s^2)^2}\}^{1/3} \tag{2.8}$$

which leads to the following equation for P-wave speed:

$$V_p = \{\frac{3E_s^2\sigma_0}{8(1-\nu_s^2)^2}\}^{1/6}(\frac{6}{\pi\rho_s})^{1/2} \tag{2.9}$$

Here, E_s, n_s, and ρ_s are respectively the elastic modulus, Poisson's ratio and density of the solid from which the spheres are made.

Using Hertz and Mindlin contact Theories, Chang et al [CMS91] computed a shear modulus of a statically isotropic packing under isotropic confining stress σ_0

$$G_{\max} = \frac{(5-4\nu)}{5(2-\nu)}\left[\frac{\sqrt{3}C_n}{\sqrt{2}\pi(1-\nu)(1+e)}\right]^{2/3}(G^2\sigma_0)^{1/3} \tag{2.10}$$

where ν and G are the Poisson's ratio and the shear modulus of the material from which the particles are made, e is the void ratio of the particulate media, and C_n is the coordination number. This shear modulus and the Poisson's ratio for the packing which was proposed as

$$\nu_{\text{pack}} = \frac{\nu}{2(5-3\nu)}$$

can be used to find a corresponding velocity-stress relation.

Micromechanics closed-form solution for G or ν of random packings subjected to anisotropic loading are not available; a limited set of numerical results were presented by Chang et al [CMS91].

Empirical Relations

Analytical models based on micromechanics suggested a power relationship between velocity and stress of particulate media. For example, equation (2.9) has a simple form

$$V_p = \alpha\sigma^{\beta} \tag{2.11}$$

where $\beta = \frac{1}{6}$ and α is a constant that depends on the mechanical parameters of the solid from which the particles are made. In this kind of micromechanics based relations, the constant exponent β depends only on the nature of contact stiffness,

e.g. $\beta = \frac{1}{6}$ for Hertzian contact, and $\beta = \frac{1}{4}$ for cone-to-plane contact [God90], while it has been observed that microfabric changes upon the increased state of stress and elasto-visco-plastic contact behavior is also important [AS95]. Considering all these factors may lead to very complicated analytical models; instead several investigators have proposed empirical expressions for this purpose.

For P-wave velocity of wave propagation in granular soil media Schmertmann [Sch78], Kopperman et al [KSK82b] and Santamarina and Potts [SP94] suggested the same power relationship of equation (2.11). For dry sand, Schmertmann [Sch78] found

$$\beta = 0.2 - 0.23$$

while Kopperman et al [KSK82b] and Santamarina and Potts [SP94] found

$$\beta = 0.22 - 0.23 \quad \text{and} \quad \beta = 0.19 - 0.22$$

respectively. For shear wave velocity, Schmertmann [Sch78] suggested the same power relationship with

$$\beta = 0.2 - 0.29$$

for dry sand. Other researchers [Roe79, KSK82a, YR84, SF95] have suggested another empirical relation to relate the shear wave velocity V_s along principal directions with the state of stress in anisotropically bounded media:

$$V_s = A\sigma_p^{\alpha}\sigma_m^{\beta} \tag{2.12}$$

Here, A, α and β are constants, σ_p is the stress in the direction of propagation, and σ_m is the stress in the direction of particle motion. Another empirical relation has been suggested by Santamarina and Cascante [SC96]:

$$V_s = A(\frac{\sigma_p + \sigma_m}{2})^{\xi}(\frac{\sigma_p - \sigma_m}{2})^{\psi} \tag{2.13}$$

where A, ξ and ψ are constants.

It must be noted that equations (2.11) to (2.13) are not valid along non-principal stress directions [AS95].

2.2.5 Wave Approach in Forward Modeling

In the wave approach of forward modeling, either elastic or acoustic wave equations are considered. The elastic wave equations for and isotropic elastic medium are

$$\rho(r)\frac{\partial^2 u_i}{\partial t^2} = (\lambda(r)+\mu(r))\frac{\partial \Delta}{\partial x_i} + \mu(r)\nabla^2 u_i, \quad i = 1,2,3 \tag{2.14}$$

$$\rho(r)\frac{\partial^2 u_i}{\partial t^2} = \mu(r)\nabla^2 \theta_i, \quad i = 1,2,3 \tag{2.15}$$

Here r is the position vector, the dilatation Δ is the volume change per unit volume, u_i represents the displacement vector components, and θ_i represents the rotation vector components. The first equation governs the P-waves and the second the S-waves. These equations can be obtained by combining Newton's law and a generalized Hooke's law [Off58].

The acoustic wave equation can be obtained by combing Newton's equation of motion with an equation of state linking changes in pressure to changes in volume [Mor48]. This leads to the following equation:

$$S(r,t) = \frac{1}{C^2(r)}\frac{\partial^2 P(r,t)}{\partial t^2} - P(r)\nabla[\frac{1}{P(r)}\nabla P(r,t)] = S(r,t) \tag{2.16}$$

Analytical solution for the elastic and acoustic wave equations to obtain the displacement field $u(r,t)$ or pressure field $P(r,t)$ as a function of space and time are available for only few simple cases [Whi83, Mor48]. Solutions of these equations for other cases are attainable using numerical methods such as the finite difference

and finite element methods. The finite difference method is used more often to obtain numerical solutions for the wave equations. Alterman and Karal [AK68], Virieux [Vir84, Vir86], Gauthier et al [GVT86], and Vafidis et al [VAK92] have used the finite difference numerical method to solve the wave equations in time domain, while Woodward and Rocca [WR88], Petrick et al [PBJ+88], and Pratt and Worthington [PW88] have used the same method in frequency-domain.

2.2.6 Reduced Wave Equation Approach of Forward Modeling

As described by Kak and Slaney [KS88], the field $u(\vec{r})$ is first considered to be the sum of two components

$$u(\vec{r}) = u_o(\vec{r}) + u_s(\vec{r}).$$

The component $u_o(\vec{r})$ is the field present without any inhomogeneities, and the component $u_s(\vec{r})$, known as the *scattered field*, will be that part of the total field that can be attributed solely to inhomogeneities. This leads to the following equation:

$$u_s(\vec{r}) = \int g(\vec{r} - \vec{r'}) o(\vec{r'}) u(\vec{r'}) d\vec{r'} \tag{2.17}$$

Here g is a known Green's function, and $o(\vec{r})u(\vec{r'})$ is the forcing function of the wave equation. Since this integral equation for the scattered field, u_s, is written in terms of the total field, $u = u_0 + u_s$, this equation cannot be solved directly, but a solution can be obtained using either the Born or the Rytov approximation.

The First Born Approximation

If we substitute $u(\vec{r}) = u_0(\vec{r}) + u_s(\vec{r})$ in equation (2.17), we obtain the following integral equation:

$$u_s(\vec{r}) = \int g(\vec{r} - \vec{r'})o(\vec{r'})u_0(\vec{r'})d\vec{r'} + \int g(\vec{r} - \vec{r'})o(\vec{r'})u_s(\vec{r'})d\vec{r'} \tag{2.18}$$

However if the scattered field, $u_s(\vec{r})$, is small compared to $u_0(\vec{r})$, the effects of the second integral can be ignored to obtain the approximation:

$$u_s(\vec{r}) \simeq \int g(\vec{r} - \vec{r'})o(\vec{r'})u_0(\vec{r'})d\vec{r'} \tag{2.19}$$

This is called the first Born approximation which is valid only when the field, $u_s(\vec{r})$, is smaller than $u_0(\vec{r})$.

The First Rytov Approximation

In the Rytov approximation which is valid under slightly different restrictions, first the total field is represented as a complex phase or

$$u(\vec{r}) = \exp \Phi(\vec{r}) \tag{2.20}$$

Here, the total complex phase, Φ, is expressed as follows:

$$\Phi(\vec{r}) = \Phi_0(\vec{r}) + \Phi_s(\vec{r}) \tag{2.21}$$

where $u_0(\vec{r}) = \exp \Phi_0(\vec{r})$ and Φ_s is the scattered complex phase. Substituting in the wave equation and simplifying leads to a differential equation that yields the following solution:

$$u_0\Phi_s = \int_{V'} g(\vec{r} - \vec{r'})u_0[(\nabla\Phi_s)^2 + o(\vec{r'})]d\vec{r'} \tag{2.22}$$

Using the Rytov approximation, we assume that the term in brackets in equation (2.22) can be approximated to $o(\vec{r'})$. This leads to the first-order Rytov approximation to the function $u_0\Phi_s$ as follows:

$$u_0\Phi_s = \int_{V'} g(\vec{r} - \vec{r'})u_0(\vec{r'})o(\vec{r'})d\vec{r'} \tag{2.23}$$

Thus Φ_s, the complex phase of the scattered field is given by

$$\Phi_s = \frac{1}{u_0(\vec{r})}\int_{V'} g(\vec{r} - \vec{r'})u_0(\vec{r'})o(\vec{r'})d\vec{r'} \tag{2.24}$$

The Rytov approximation is valid under a less restrictive set of conditions than the Born approximation [KS88]. Both of these approximations are the basis of the "Fourier Diffraction theorem" which is the base of diffraction tomography.

2.2.7 Ray Based Approximation of Forward Modeling

Born and Wolf [BW65] used the wave equation to show that when the wavelength of the propagation energy is small compared to the size of the scatterers, or when the changes in the transmission speed per unit wavelength are small, energy may be regarded as being transported along curves that are orthogonal trajectories to the moving wave fronts.

The Eikonal Equation

The basic equation of geometrical acoustics describing the propagation of wave fronts in an inhomogeneous medium of a continuously varying wave propagation speed, c, is the eikonal equation

$$|\vec{\nabla}\Theta|^2 = \frac{\rho}{k} = \frac{1}{c^2} \tag{2.25}$$

where the surfaces $\Theta(r)$ = constant may be called the geometrical wave fronts, and ρ and k are the density and bulk modulus of the gas or fluid through which the wave is propagating. As described by Nolet [Nol87], this equation can be obtained by applying Fourier transform to the acoustic equation and then assuming that the frequency $\omega \to \infty$. The vector $\vec{\nabla}\Theta$ is perpendicular to the wave front, and therefore by definition parallel to the ray. Letting $d\vec{r}$ be a tangent along the ray with length ds, equation (2.25) leads to the following second-order differential equation for rays:

$$\nabla(\frac{1}{c}) = \frac{d}{ds}(\frac{1}{c}\frac{d\vec{r}}{ds}) \tag{2.26}$$

Beginning with the equation of elastic waves and following a similar way, we can obtain the following eikonal equations of wave propagation in isotropic elastic media:

$$|\vec{\nabla}\Theta|^2 = \frac{\rho}{\lambda + 2\mu} = \frac{1}{V_p^2} \tag{2.27}$$

$$|\vec{\nabla}\Theta|^2 = \frac{\rho}{\mu} = \frac{1}{V_s^2} \tag{2.28}$$

Here ρ is the solid density, λ and μ are Lame's elastic constants, and V_p and V_s are the P-wave and S-wave velocities.

The Fermat's Principle in Ray Tracing

Fermat's principle is an important and very useful principle that was formulated in the $17th$ century originally for optical rays. This principle states that energy propagates along the path that makes travel-time minimum. If we consider the energy as propagating along a generally curved ray between two points through a medium with $V(\vec{r})$ representing the speed of energy transfer, the following integral equation will relate the travel-time t to $V(\vec{r})$:

$$t = \int_R \frac{dr}{V(\vec{r})} \tag{2.29}$$

where R represents the curved ray and dr is an infinitesimal portion of the ray in the neighborhood of r.

Another way to obtain the eikonal equations, as shown by Stavroudis [Sta72] and Nolet [Nol87], is to use Euler's equation in "Calculus of Variation" to minimize the travel-time in equation (2.29).

Validity of the Ray Approximation

Many researchers have studied the influence of diffracted waves when the ray approximation is incorporated. Wielandt [Wie87] concluded that ray approximation works well for positive anomalies, but in the presence of a negative anomaly, travel-time data are likely to present the travel-time of the fastest path around the anomaly, even at relatively short distances. In cases where diffracted waves have noticeable amplitudes, if we use ray based models, inclusions with a positive anomaly will appear larger than they are in nature; the amplitude of negative anomalies may be grossly underestimated.

Methods of Ray Tracing

There are basically two main methods for tracing rays between two given end-points through a continuously varying field: initial value and boundary value methods. In the first method, often called the *shooting method*, we shoot the ray from one end-point in a certain initial direction, find its destination, and then systematically vary the initial direction until the ray emerges at the desired end-point. In the second method, often called the *bending method*, we vary the curved path between the end-points until it satisfies Fermat's principle. The shooting methods are simple and straightforward, but on the other hand much more time consuming than bending

methods. According to Julian and Gubbins [JG77], they sometimes take 10 times as much as bending methods. The bending methods, on the other hand, are fast and efficient.

Shooting Methods

Wesson [Wes71] developed a numerical technique to trace seismic rays in an isotropic-inhomogeneous medium. In that work, he used the central finite difference approximations to the ray equation.

Julian and Gubbins [JG77] considered another formulation for the ray equation in the 3-D case which leads to an initial value problem for a system of first-order differential equations. To solve this system, the starting direction corresponding to the desired end-point is found iteratively.

Johnson et al [JGSJ75] considered the ray equation (2.26) as the basis for a ray tracing algorithm. This equation, in combination with a truncated Taylor series, leads to the following expression suitable for numerical implementation:

$$r(s+\Delta s) = r(s) + \frac{dr}{ds}\Delta s + \frac{1}{2n}[\Delta n - (\Delta n \frac{dr}{ds})\frac{dr}{ds}](\Delta s)^2 \qquad (2.30)$$

where n is the refraction index. One systematic way for finding the proper starting direction in the shooting methods is using Newton-Raphson method as used by many researchers ([LD80, AK82]).

The method of characteristics is based on a direct solution to the eikonal equation. This method was initially introduced by Jakowatz and Kak [JK76]; for the 2-D case it leads to the following numerical algorithm:

$$\begin{aligned} x_{j+1} &= x_j + (\tfrac{p_j}{n})\Delta s \\ y_{j+1} &= y_j + (\tfrac{q_j}{n})\Delta s \\ p_{j+1} &= p_j + \tfrac{\partial n}{\partial x}\Delta s \\ q_{j+1} &= q_j + \tfrac{\partial n}{\partial y}\Delta s \end{aligned}$$

where $p = \frac{\partial s}{\partial x}$ and $q = \frac{\partial s}{\partial y}$.

Lytle and Dines [LD80] introduced a numerical algorithm which is based upon a Runge-Kutta solution of the 2-D ray equation:

$$\frac{d\Theta}{ds} = \frac{1}{n}(\cos\Theta\frac{\partial n}{\partial y} - \sin\Theta\frac{\partial n}{\partial x}) \tag{2.31}$$

They used a Newton-Raphson method to obtain the correct starting Θ.

White [Whi89] and Bregman et al [BBC89a, BCB89] used a velocity field with piecewise constant gradient in a mesh of triangular cells. They proved the analytical ray path in each cell is a circular arc, thus their ray tracing method is a matter of connecting these circular arcs starting by a proper shooting direction which is found iteratively.

Bending Methods

Wesson [Wes71] used the central difference approximations to the ray equation to formulate the bending method for the seismic-ray tracing problem. Chander [Cha75] developed an equivalent method by approximating the integral equation of the travel-time (2.29) by a sum, and then solving for the minimum time-path directly.

Julian and Gubbins [JG77] developed a method which differs slightly from those of Wesson [Wes71] and Chander [Cha75]; they expressed the ray differential equa-

tions in terms of changes in the ray-path, linearized, and solved them iteratively after applying the finite difference approximations.

Yang and Lee [YL76] introduced another approach based on bending; they first reduced the second-order ray expressions to a set of first-order equations and then solved them by applying a finite difference technique. This approximation leads to a system of nonlinear equations that must be solved numerically. Pereyra et al [PLK80] followed the same approach and applied the Newton-Raphson method to solve the system of nonlinear equations.

To avoid falling in local minima, Thurber and Ellsworth [TE80] proposed approximating the complex heterogeneous medium, by a layered medium by laterally averaging velocities in region local to the source and receivers. Ray-paths in the approximated model can be easily determined and used as the starting paths. For more accurate ray tracing routines, Um and Turber [UT87] proposed an approximate fast algorithm as a bending method. Rather than explicitly solving the ray expression, they perturbed a starting path to minimize the travel-time. The perturbation is made according to a geometrical interpretation of the ray equation. Prothero et al [PTE88] developed another ray tracing fast algorithm. They first determined the best circular arc path between two end-points which minimizes the travel-time, and then made perturbation consisting of a sum of N sinusoidal terms of the circular arc path to minimize the travel-time. To find the amplitudes of the sinusoidal terms which minimizes the travel-time, they used the simplex method of Nelder and Mead [NM65]. Berryman [Ber90] applied a similar algorithm in which a straight line rather than a circular arc path is perturbed by sinusoidal terms.

Vidale [Vid88] suggested a ray tracing method based on solving the eikonal equation (2.25). That method appeared to be accurate. The algorithm includes two steps: In the first step, travel-time values at all grid-points are calculated by

using a finite difference approximation to the eikonal equation. At the second step, ray-paths with minimum travel-times are located by tracing the path of steepest gradient of travel-time from the receivers back to the source. This type of method is often called a forward-and-backward method. Asakawa and Kawanaka [AK93] proposed another ray tracing method based on the forward-backward method. In this method, they divided the region into a number of cells, and considered an extra number of calculation points on each cell boundary. They assumed that the travel-time at any point between two other calculation points on the boundary can be linearly interpolated in terms of the travel-times of the calculation points. In the forward process, the travel-times of the calculation points are computed successively using the previously calculated travel-times on the points in the neighborhood. In the backward process, they used the procedure in [Vid88].

Ray Tracing in Anisotropic Media

Based on the general form of the elastic wave equation

$$\frac{\partial}{\partial x_i}(c_{ijkl}\frac{\partial U_k}{\partial x_j}) = \rho\frac{\partial^2 U}{\partial t^2} \tag{2.32}$$

Červený [Č72] derived the eikonal equations which govern compressional waves and shear waves propagating in heterogeneous-anisotropic media. To derive these equations, he assumed elastic constants that are continuous functions of position. Those equations were derived using the theory of ray series and they lead to a system of ordinary first-order differential equations to be solved for obtaining ray paths. That method of ray tracing appeared to be cumbersome; it is impractical to apply for iterative inversion algorithms where ray tracing must be iteratively carried out.

Tanimoto [Tan87] presented ray expressions for surface waves in anisotropic laterally inhomogeneous media. Those equations are derived in two ways: first from the Hamilton's canonical equations, and second from Fermat's principle.

2.3 Closure

In this chapter, we have presented a literature review of methods of forward modeling in computerized tomography. Here, we will summarize each presented method.

In section 2.2.2, we addressed the issue of anisotropy in granular soil. In that section, we pointed out that there are two kinds of soil anisotropy: inherent anisotropy and stress induced anisotropy. A considerable amount of research has been done in relation to wave propagation in cross-anisotropic media. The outcome of this research is some analytical relations for calculating phase and ray velocities. We have not come across any research work which addresses wave propagation in stress induced anisotropic media.

In section 2.2.4, we first reviewed analytical velocity-stress relations in particulate media based on micromechanics. Those relations are somewhat complicated, and can be applied in only limited circumstances. On the other hand, empirical relations are less complicated and are widely used in geotechnical applications.

In section 2.2.5, we discussed various wave approaches for forward problems. These approaches are based on the p.d.e's of the wave propagation which are, theoretically, capable of providing high-resolution images. They require a huge amount of computation, and because of high nonlinearity involved in the governing equations, numerical algorithms may lead to local minima. For stress induced anisotropic media, the governing p.d.e's of wave propagation are not available, so

that this approach is not applicable.

The reduced wave equation approach which is reviewed in section 2.2.6 is based on a simplified version of the wave equations, and is applicable when the scattered field is small. This approach leads to a forward process which is more efficient than the wave approach; because the Green's function is not available for stress induced anisotropic media, it is not applicable in this study.

In section 2.2.7, we reviewed ray based methods of forward modeling. These methods can be applied when the wavelength is small compared to the size of scatterers, or when the inhomogeneity is continuous. The most important part of the forward modeling is ray tracing, for which two sets of methods are reviewed: shooting and bending. The shooting methods are simple and straight forward, but time consuming compared to bending methods. Bending methods are fast and efficient, and can be divided into two sets: those based on the ray differential equation, and those based on perturbation methods. Perturbation based methods are widely used and have proved to be successful in many applications. For ray tracing in cross-anisotropic media, the eikonal equation has been derived and used for ray tracing. For ray tracing in stress induced anisotropic media, to our knowledge, no method has been suggested; perturbation methods used in isotropic media can be easily extended to this kind of anisotropic media. In chapter 4, we will present a ray tracing technique for stress induced anisotropic media based on perturbation methods.

Chapter 3

Inverse Problem in CT

3.1 Introduction

The inverse problem, in general sense, is the determination of a parameter or a set of parameters, e.g., velocity of wave propagation from a set of measured parameters, e.g., travel-time. In tomography, the spatial distribution of parameters is of interest, therefore the purpose of solving inverse problems arise in tomography applications is the determination of the spatial distribution of some parameter(s) from a set of measured parameter(s). In this chapter, we will present a review on techniques and schemes commonly used in the inverse problem of geophysical tomography.

3.2 Inverse Problem in the Wave Approach

In Section 2.2.5 of Chapter 2, the wave approach was presented for forward modeling. This approach is based on the solution of the wave differential equations. One of the earliest attempts to apply the wave equations in inverse problems was made

by Bleinstein and Cohen [BC79], and Cohen and Bleinstein [CB79]. In that work, the acoustic wave equation is simplified in such a way that the inversion of reflected data leads to a closed form solution. The approximation used in the simplified wave equation is valid for small transverse variations compared to vertical variations.

In Tarantola [Tar84a] and Gauthier et al [GVT86], the exact form of the acoustic equation is used for the forward modeling. In that work, a nonlinear inversion procedure using a least-squares method has been used to minimize the misfit of the calculated and observed pressure field $P_0(r,t)$ for discrete points on the surface. This leads to minimization of a nonquadratic function in a multi-variable space. The procedure uses the gradient method in optimization; forward modeling, as a part of the inversion procedure, is based on a finite difference approximation of the acoustic wave equation in the time domain.

Tarantola [Tar84b] and Mora [Mor87] have tackled the inverse problem of the elastic wave. Following Tarantola [Tar84a], they have used the same nonlinear inversion strategy to obtain the solution which minimizes the misfit of calculated and observed displacements $u(r,t)$ at discrete points on the boundary.

The strategy presented by Tarantola [Tar84a, Tar84b], Gauthier et al [GVT86], and Mora [Mor87] which is based on the wave equations in time domain has some serious drawbacks, e.g., consuming huge computational time. To overcome this drawback, Pratt and Worthington [PW90] and Pratt [Pra90] carried out the forward modeling procedure in the frequency domain rather than in the time domain. For the forward problem part, they applied the finite difference technique to solve the wave equations in the frequency domain; for inversion, they applied the strategy used by Tarantola [Tar84a].

Because of the highly nonlinear relation between data and the desired model,

and on the other hand the large number of unknowns involved in the wave approach, traditional optimization techniques are likely to converge to a local minimum, unless the starting point is adequately close to the global minimum. To provide such a starting point, Pratt and Goutly [PG91] suggested travel-time tomography, which is based on ray approximation as a proper way for obtaining acceptable low resolution images. In that work, they used the same strategy used by Pratt and Worthington [PW90] and Pratt [Pra90].

In Tarantola [Tar84a] and following work, the entire output wave pattern is chosen to construct the misfit function; this is the main reason of the highly nonlinear relation between the data and model space. To overcome this problem, Lou and Schuster [LS91] have moved in another direction; they reduced the nonlinearity by expressing the wave equation in terms of travel-time, thus choosing the travel-time parameter to construct the misfit function rather than the entire output wave pattern.

3.3 Transform Based Inversion Methods

Using Born or Rytov approximations, the Fourier diffraction theorem can be obtained [KS88]: When an object is illuminated by a plane wave, the Fourier transform of the forward scattered field measured on a line parallel to the plane wave gives the values of the 2-D Fourier transform of the object along a semicircular arc in the frequency domain. As the frequency of wave increases, the radius of arc increases until the scattered field is closely approximated by the Fourier Slice Theorem which is stated as [Kak85]: The Fourier transform of a parallel projection of an image $f(x, y)$ taken at an angle Θ gives a slice of the two-dimensional transform $F(u, v)$, subtending the angle Θ with the u-axis.

To apply these theorems in tomography, we must illuminate the object by plane waves from many directions. In the way we may obtain the two-dimensional Fourier transform of the object and finally recover the object by Fourier inversion.

Devaney [Dev82] used the Rytov approximation and developed formulas that operate on the monofrequency scattered field. Devaney [Dev84] discussed geophysical applications of diffraction tomography.

One of the major problems in geophysical applications of the Fourier diffraction theorem is that the region cannot be imaged by parallel beams because such plane wave–forms cannot be generated exactly. To overcome this problem, Whitten and Molyneux [WM88] have presented an extension of the algorithm suggested by Devaney [Dev82], which allows for the use of sources with arbitrary radiation patterns.

Extraction of the scattered wave field from the total recorded wave field is a serious difficulty for the implementation of diffraction tomography. This problem reduces to the problem of estimating the source signature. Pratt and Worthington [PW88] have shown how the source function can be extracted statically from the data.

3.4 Ray Based Travel-Time Tomography

By introducing $S(\vec{r}) = \frac{1}{V(\vec{r})}$, the slowness of wave propagation, we may use equation (2.29) to obtain the travel-time equation:

$$t = \int_R S(\vec{r}) dr \tag{3.1}$$

This equation is a Fredholm integral equation of the first kind [dH80] which can be written down for each pair of source-receiver on the boundary of the region. The

purpose of inversion is to determine $S(\vec{r})$, the slowness distribution of the region of interest, from a given set of measured travel-times t.

If the ray paths are known, then by discretizing the region into small cells and supposing that the ray path in the cell is a straight line along which the speed is constant, we can obtain the following linear algebraic equations:

$$\sum_j \Delta r_{ij} s_j = t_i \tag{3.2}$$

This is the discretized version of (3.1); here t_i is the measured travel-time of the ith ray and Δr_{ij} is the length of the portion of the ith ray contained in the jth cell.

Combining the equations for all the rays, we obtain the matrix equation

$$[M]\{S\} = \{t\} \tag{3.3}$$

where $[M]$ is a matrix that depends on the ray path and $\{S\}$ and $\{t\}$ are the vectors of slowness and the measured travel-times, respectively.

Many researchers (e.g., [DL79, McM83, Iva87]) considered the rays as straight lines; under this assumption, the problem becomes linear and, consequently, relatively easy to invert. However, in most applications, it is well-known that ray bending is significant where spatial variations of speed are considerable. Thus, neglecting the ray-bending in such regions is not reasonable, and the straight line assumption is valid only in limited cases.

3.4.1 Reconstruction Techniques in Straight-Ray Tomography

Even though the straight ray assumption in geophysical applications is valid in only very limited cases, some researchers (e.g. [LL77, DL79, McM83, Iva87]) have used

reconstruction techniques that are based on this assumption. These techniques are essentially based on solving the system of equations (3.3); this can be carried out using iterative methods or matrix methods.

Iterative Methods

In tomographic imaging applications, iterative methods based on the "method of projections" have been applied successfully. This method, first proposed by Kaczmarz [Kac37] and elucidated by Tanabe [Tan71], is based on the system of linear equation as follows:

$$\sum_{j=1}^{N} m_{ij} S_j = t_i \quad i = 1, 2, \cdots, M \tag{3.4}$$

where M and N are the number of equations and unknowns, respectively. Each of these equations represents a hyperplane in N-dimensional space, and the image may be considered as a single point, the intersection of the hyperplanes in this space. The computational procedure for locating this point consists of first starting with an initial guess on the first hyperplane, projecting this point on the next hyperplane successively and finally projecting back onto the first hyperplane, and so forth. If a unique solution exists, the iterations will always converge to that point; however, if the system is singular, the iterations converge to a point with minimum distance to the initial guess point [Tan71]. The process of projecting point $S^{(i-1)}$, obtained from the $(i-1)$th iteration on the hyperplane represented by the ith equation to yield point $S(i)$ can be mathematically described by:

$$S^{(i)} = S^{(i-1)} + \frac{t_i - \overline{t_i}}{\sum_{k=1}^{N} m_{ik}^2} m_{ik} \tag{3.5}$$

where t_i is the measured travel-time along the ith ray, while $\overline{t_i}$ is the computed travel-time based on the $(i-1)$th solution. This direct implementation of the

method for a system with a large number of views with a large number of unknowns is expensive, but there are many approximate algorithms based on the idea: ART, SIRT, SART, etc.

In ART (Algebraic Reconstruction Technique), in order to speed the computation, the m_{ik}'s in the last equation are simply replaced by 1 or 0, depending upon whether the center of the kth image cell is or is not within the ith ray [KS88], and the values of S are corrected after each iteration. Noise in the measurements, and the procedure used for correction after each iteration, lead to poor images. This technique has been used for geotomography by Larger and Lytle [LL77] and MacMechan [McM83].

SIRT (Simultaneous Iterative Technique) uses the same approximations as in ART, but applied the correction only after all available rays have been examined. This approach, though slower than ART, usually leads to better results [KS88]. Dines and Lytle [DL79] applied this technique for imaging attenuation and speed in crosshole tomography. They applied SIRT in combination with averaging scheme for smoothing.

SART (Simultaneous Algebraic Reconstruction Technique), first reported in Anderson and Kak [AK84], is another method that combines the best of ART and SIRT to get reconstructions of good quality and numerical accuracy in only one iteration [KS88]. In this method, as in SIRT, the correction terms are simultaneously applied for all the rays in one projection.

Matrix Methods

To invert the linear systems that appear in the computational process in tomographic applications, matrix methods are widely applied. To employ the existing

techniques for inversion purposes, there are some difficulties that should be considered. The most important difficulty is that the matrix of coefficients that appears in geotomographic applications is neither square nor of full rank, and also the systems are mixed-determined. The other aspect of the linear systems is that the vector of measured data is not accurate because of the inevitable noise. Singular value decomposition can be employed to analyze the effect of noise on the solution of equation (3.3).

According to Lancsoz [Lan61], any $M \times N$ matrix G can be written as the product of three matrices as follows:

$$G = U\Lambda V^T \tag{3.6}$$

where U and V are orthogonal matrices of dimensions $M \times M$ and $N \times N$ respectively, and Λ is an $M \times N$ diagonal matrix whose diagonal elements are called "singular values". The generalized inverse of G denoted by G^+ including only the r non-zero singular values is:

$$G^+ = V_r \Lambda_r^+ U_r^T \tag{3.7}$$

Therefore, applying the so-called natural solution of any linear system such as $Gx = d$ can be expressed as:

$$\tilde{x} = G^+ d \tag{3.8}$$

This natural solution reduces to the minimum length solution when the problem is purely under-determined, and to the least squares solution when the problem is purely over-determined [Men89]. An equivalent expression for the natural solution is:

$$\tilde{x} = \sum_{i=1}^{r} \frac{\alpha_i}{\lambda_i} V_i \tag{3.9}$$

where α_i are coefficients related to the measured data, and λ_i and V_i are the singular values and the columns of V respectively. In geophysical applications, there are

commonly a number of small singular values among the non-zero ones; these cause amplification of the noise carried by α_i; SVD must be modified.

Damping stabilizes the natural solution when the data contains noise; equation (3.9) is replaced by:

$$\tilde{x} = \sum_{i=1}^{r} \frac{\lambda_i \alpha_i}{\lambda_i^2 + \mu^2} V_i \tag{3.10}$$

When λ_i is small, the contribution of the corresponding eigen vector is small so that the noise is not amplified; when λ_i is large compared to μ, the latter has little effect. The natural solution can be written:

$$\tilde{x} = V_r \{\Lambda_r (\Lambda_r^2 + \mu^2 I)^{-1}\} U_r^T d \tag{3.11}$$

Another way to obtain equation (3.11) is to minimize a combination of solution error and length ($\Phi = \varepsilon^T \varepsilon + \mu^2 (x^T x)$), where $\varepsilon = Gx - d$. This minimization leads to:

$$\tilde{x} = (G^T G + \mu^2 I)^{-1} G^T d \tag{3.12}$$

This is called the Ridge regression estimate or the damped-least-squares solution [Men89], and it has been applied by many researchers such as Wiggins [Wig72] who used equation (3.11), Nolet [Nol85] who compared some inversion algorithms used to obtain the solution, Bregman et al [BBC89a, BCB89, BBC89b], White [Whi89], Berryman [Ber90] and Phillips and Fehler [PF91]. Berryman [Ber90] applied another version of the damped-least-squares solution which is based on minimizing a combination of solution error and the distance between the solution and another starting point.

The damped least-squares method can be generalized, if a combination of solution error and some functional such as Cx rather than solution length is minimized. Thus, ($\|Gx - d\|^2 + \mu^2 \|Cx\|^2$) can be minimized to yield the following regularization

equation:

$$(G^T G + \mu^2 C^T C)x = G^T d \tag{3.13}$$

The damped least-squares solution is a special case where $C = I$. In practice C is chosen to smooth out the oscillations in the solution caused by noise; often Cx is taken so that it approximates the second derivative of the solution. Equation (3.13) has been used by many researchers such as Lytle and Dines [LD80] who applied this technique using the Laplacian operator to construct C in borehole to borehole tomography, and Ivansson [Iva87].

Equation (3.9) shows that $\tilde{x}$ is linearly dependent on the eigen- values of the G matrix. This means that errors in G arising from modeling approximations will lead to errors in $\tilde{x}$; these are compounded with errors arising from noisy data. The total least-squares solution attempts to find a solution that minimizes these errors in some sense. Golub and Van Loan [GVL80] presented a singular value decomposition analysis of this problem, and Weldon Demmel [Dem87], Van Huffal and Van dewalle [VHV88] have modified their algorithm.

3.4.2 Iterative Schemes in Curved-Ray Tomography

Within the last several years, researchers in tomography have recognized that in the case of strongly refracting media, straight-line inversion can lead to significant reconstruction error. This has lead to the development of reconstruction techniques in curved-ray tomography.

In curved-ray tomography, we attempt to determine $S(\vec{r})$, the slowness distribution from a set of measured travel-times that are governed by equation (3.1). In this equation, R is governed by the Eikonal equation. In general, the problems defined by equation (3.1) and the Eikonal equation are highly nonlinear, therefore

iterative schemes suggested to solve them are not always guaranteed to converge to desired solutions.

Straight-Line Inversion with Modified Data

Straight-Line Inversion with Modified Data (SLIM) is an iterative approach that involves the straight-line inversion of a modified data set that is updated in each iteration. The measured data, e.g., t are used to produce an initial field approximation, e.g. $S^{(0)}$, and then for each subsequent iteration k, the data are adjusted by the difference between the data $\tilde{t}^{(k-1)}$ predicted on the basis of preceding iterate $S^{(k-1)}$ and the true data set t. That is, each iteration is characterized by the straight line inversion of modified data $\bar{t}$, where

$$\bar{t}^{(k)} = \bar{t}^{(k-1)} + (t - \tilde{t}^{(k-1)}) \tag{3.14}$$

where $\bar{t}^{(0)} = t$. The predicted data $\tilde{t}^{(k)}$ are normally produced by applying well established ray tracing methods, and iteration continues until some termination criterion for either the predicted data set $\tilde{t}^{(k)}$ or the field iterations $S^{(k)}$ is satisfied.

The SLIM process, which is illustrated in Figure 3.1, has been used by researchers in different tomography disciplines; Ivansson [Iva85] used this technique in geophysical imaging.

Curved Ray Algebraic Inversion

In Curved Ray Algebraic Inversion (CRAI), straight-line inversion is used only to provide an initial iterate $S^{(0)}$. Subsequent approximations $S^{(k)}$ are obtained by curved-ray inversion where the ray paths are defined on the basis of the preceding iterate $S^{(k-1)}$, and this inversion process typically involves discretization of the

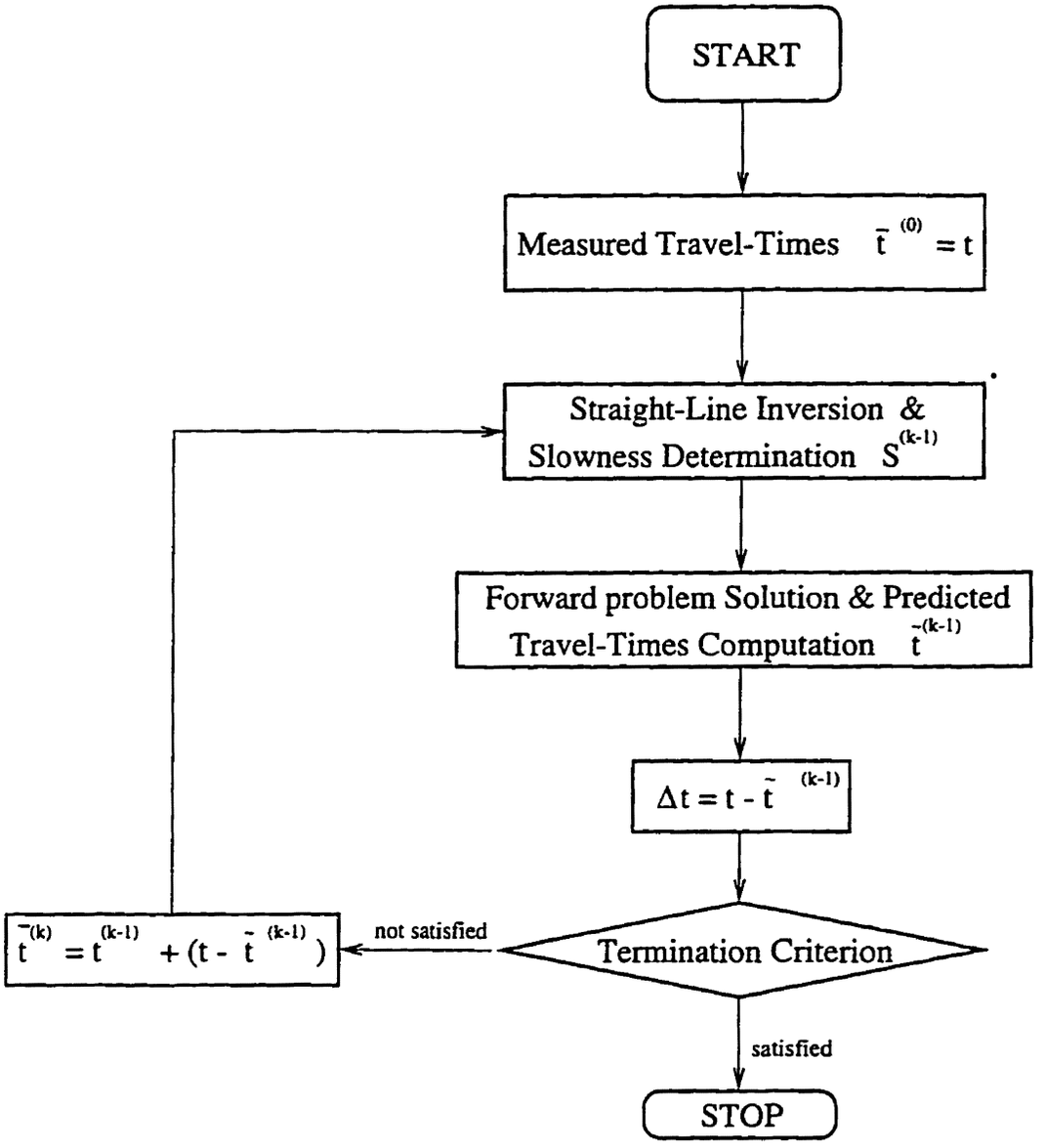


Figure 3.1: SLIM-Straight line inversion with modified data

field followed by solution of the resulting system of algebraic equations. The CRAI concept which is outlined in Figure 3.2 has been applied by many researchers such as Lytle and Dines [LD80] who applied it in borehole to borehole tomography.

Although the use of straight-line inversion leads to an algorithm that is faster and easier to implement than CRAI, SLIM appears to be a rather unsophisticated method where the data are adjusted to counteract the inaccuracy due to the conventional inversion approach used.

Generalized Iterative Schemes in Curved-Ray Tomography

Dolovich [Dol90] showed that both SLIM and CRAI can be interpreted as mappings of the general form:

$$S = S + O^{-1}(t - MS) \tag{3.15}$$

where M is the operator which maps S, the slowness vector, to t, the travel-times vector, and O^{-1} is an inverse operator. Using the contraction mapping theorem, Dolovich has studied and analyzed the convergence properties of mappings of the form given by equation (3.15) in general and of SLIM and CRAI in particular. He finally introduced the idea of "generalized curved ray algebraic inversion" based on equation (3.15), the proposed methods to select the inverse operator O^{-1} so as to enhance the convergence properties of the iterative scheme corresponding to this equation.

Perturbation Method

In this method, a linear equation is obtained by approximately relating the slowness model modifications to the travel-times perturbations. Having an estimate of the

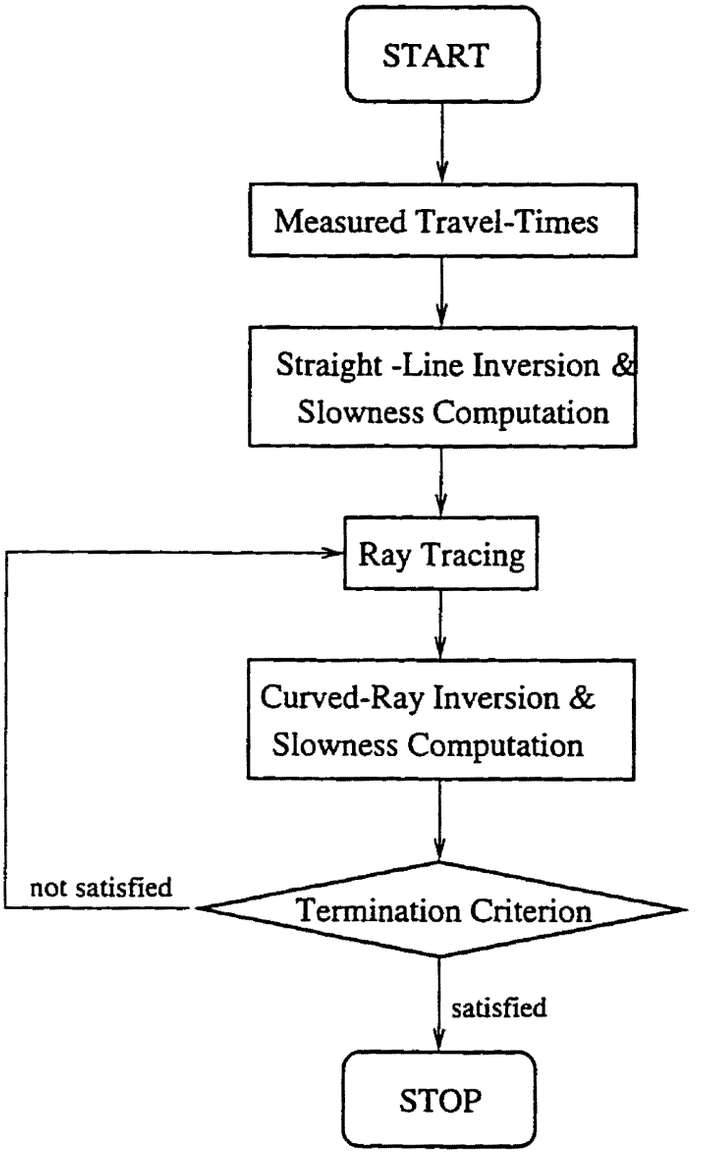


Figure 3.2: CRAI-Curved ray algebraic inversion

desired model, $\{S_0\}$, and considering the basic travel-time equation (3.1), we can obtain a relation in two ways.

If we consider R_0 as a representative of rays that correspond to $\{S_0\}$, we will have a typical equation:

$$\int_{R_0} S_0(\vec{r})dr = t_0 \tag{3.16}$$

where t_0 represents the calculated travel-times using the $\{S_0\}$ model. The matrix version of this equation is:

$$[M]\{S_0\} = \{t_0\} \tag{3.17}$$

Fermat's principle implies that first-order variations of the slowness will cause second-order perturbations in ray paths [AR80], therefore we can consider the same $[M]$ as the approximated matrix corresponding to the true model $\{S\}$; thus

$$[M]\{S\} = \{t\} \tag{3.18}$$

where $\{t\}$ is the vector of the measured travel-times. Subtracting the last two equations from each other, we obtain the following linear equation:

$$[M]\{\Delta S\} = \{\Delta t\} \tag{3.19}$$

where $\{\Delta S\}$ and $\{\Delta t\}$ are the vector of slowness modifications and the vector of travel-time perturbations respectively. Having found $\{\Delta S\}$ from equation (3.19), we can find a new estimate of slowness $\{S_0 + \Delta S\}$ and use it to construct a new $[M]$ matrix, and so on so forth.

This technique has been used by many researchers such as Nolet [Nol85], Bording et al [BGL+87], and Michelena and Harris [MH91].

Another formulation can be obtained by considering the basic equation of travel-time tomography in terms of wave speed:

$$t[V(r)] = \int_R \frac{dr}{V(r)} \tag{3.20}$$

where $t[V(r)]$ is the observed travel-times as a function of wave speed. To linearize this equation, we consider a Taylor expansion about a starting speed model $V_0(r)$:

$$\begin{aligned} t[V(r)] = \quad & t[V_0(r)] + \frac{\partial t}{\partial V}\Big|_{V=V_0} \delta V(r) \\ & + \frac{\partial t}{\partial R}\Big|_{V=V_0} \frac{\partial R}{\partial V}\Big|_{V=V_0} \delta V(r) + O(\delta^2 V(r)) \end{aligned} \tag{3.21}$$

According to Fermat's principle, if V_0 is close enough to the actual V distribution, then

$$\frac{\partial t}{\partial R}\Big|_{V=V_0} \cong 0$$

Thus, discarding terms of $O(\delta^2 V(r))$, we obtain the following linear equation:

$$\delta t = -\int_{R_0} \frac{\delta V(r)}{V_0^2(r)} dr \tag{3.22}$$

This linear relation between δt and $\delta V(r)$ has been used by various researchers such as Julian and Anderson [JA68], Backus and Gilbert [BG69], Bregman et al [BBC89a, BCB89, BBC89b] and White [Whi89].

Feasibility Concept Based Method

This method, introduced by Berryman [Ber89b, Ber89a, Ber90, Ber92] as a stable iterative technique in nonlinear travel-time tomography, can be applied in various tomography problems [Ber92]. Fermat's principle shows that trial wave speed models that lead to ray paths with travel-times smaller than the observed travel-times are not feasible. He introduces a feasibility violations number, which is the number of ray paths with travel-time less than that measured, and applies Fermat's principle to show that the feasible region is convex. The other important idea which is used repeatedly in the algorithm is that the ray paths which correspond to a given slowness model $\{S\}$ are valid for any other slowness model $\{\gamma S\}$ where $\gamma > 0$.

These ideas constitute the basis for the following iterative algorithm that has been applied successfully in travel-time tomography problems [Ber89b, Ber89a]:

1. Chose an initial estimate of the slowness model $\{S\}$, and form the system of linear equations for ray tracing.
2. Compute γ such that $\{\hat{S}\} = \{\gamma S\}$ best fits the observed travel-times for these rays.
3. Compute $\{S_\mu\}$, the modified model of $\{S\}$ which gives a better fit to the data using the damped least-squares method.
4. Compute a model between $\{\hat{S}\}$ and $\{S_\mu\}$ which is closer to the boundary of the feasible region. Since the feasible region is convex, such a model exists.
5. Scale $\{S_m\}$ by a proper γ to find $\{S_f\} = \{\gamma S_m\}$ which minimizes the number of feasible violations.
6. Stop if $\{S_f\}$ properly fits the data, otherwise repeat steps 2 to 6.

3.4.3 Ray Based Travel-Time Tomography in Inhomogeneous Anisotropic Media

The inverse problem of travel-time tomography of inhomogeneous-anisotropic media is essentially the same as that of inhomogeneous-isotropic media. The inversion process basically consists of two components: a ray tracing algorithm, which is the forward problem solver in inhomogeneous-anisotropic media, and an iterative scheme. For ray tracing, the method proposed by Červený [Č72] can be used; for the second component, one of the iterative schemes described earlier can be

used. However, there are some difficulties that arise in the inversion process of inhomogeneous-anisotropic media in particular. One difficulty is that the number of unknowns in inhomogeneous-anisotropic media is much larger that that in inhomogeneous-isotropic media; this leads to problems that are considerably more ill-conditioned than those that arise in the isotropic case. The other difficulty is the long time required to solve the forward problem. These difficulties must be dealt with properly, otherwise the inversion process is impractical.

Červený and Jech [CJ82] proposed a linearized solution to evaluate travel-times of seismic body waves in inhomogeneous, slightly anisotropic media. In this method, they make a correction to the travel-time according to the unperturbed ray paths which were calculated according to a given C_{ijkl}. Based on this work, Červený and Firbas [CF84] suggested a numerical model for travel-time inversion in inhomogeneous-anisotropic media.

Chapman and Pratt [CP92, PC92] presented a travel-time tomography technique for inhomogeneous, slightly anisotropic media. They basically followed Červený [$\check{C}$72] in their formulation, however they used an isotropic ray tracing scheme, reckoning that anisotropic effects are due to weak perturbation of the isotropic system, and thus can be ignored in ray tracing. For the inversion process, they used the perturbation method which was described in the Chapter 2.

3.5 The Nonlinear Least-Squares Problem

The nonlinear least-squares problem arises most commonly from data fitting applications, where the purpose is to fit a set of data such as travel-time data with a model such as $t(a)$ that is nonlinear in a. Such a problem arises in the research

and will be discussed in the coming chapters. This problem is stated as:

$$\text{Minimize } f(x) = \frac{1}{2}R(x)^T R(x) = \frac{1}{2}\sum_{i=1}^{m} r_i(x)^2 \tag{3.23}$$

where $x^T =< x_1, x_2, \cdots, x_n >$ represents the vector of unknowns and $R^T(x) =< r_1(x_1), r_2(x_2), \cdots; r_n(x_n) >$ where $r_i(x)$ denotes the ith residual, represents the vector of residual functions. It is usually not recommended to solve nonlinear least-squares problems by general-purpose minimization methods; this is mainly due to the special structure of equation (3.23) which leads to special relations between the derivatives of $R(x)$ and $F(x)$.

The first derivative matrix of $R(x)$ is simply the Jacobian matrix $J(x)$ defined as:

$$J_{ij}(x) = \frac{\partial r_i(x)}{\partial x_j}, i = 1, \cdots, m,\ j = 1, \cdots, n$$

Thus, an approximated model of $R(x)$ around a point x_c is:

$$M_c(x) = R_c(x) + J(x_c)(x - x_c) \tag{3.24}$$

On the other hand, it can be shown that the first and second derivatives of $f(x) = \frac{1}{2}R(x)^T R(x)$ are:

$$\nabla f(x) = \frac{\partial f}{\partial x_j} = J(x)^T R(x), j = 1, \cdots n \tag{3.25}$$

and

$$\nabla^2 f(x) = \frac{\partial^2 f}{\partial x_i \partial x_j} = J(x)^T J(x) + S(x), i = 1, \cdots n,\ j = 1, \cdots n \tag{3.26}$$

where

$$S(x) = \sum_{i=1} m r_i(x) \nabla^2 r_i(x)$$

Thus the quadratic model of $f(x)$ around x_c is:

$$m_c(x) = \frac{1}{2}R(x_c)^T R(x_c) + R(x_c)^T J(x_c)(x - x_c) + \frac{1}{2}(x - x_c)^T (J(x_c)^T J(x_c) + S(x_c))(x - x_c) \tag{3.27}$$

To obtain the minimum of $m_c(x)$, we must have $\nabla m_c(x_+) = 0$; this leads to Newton's method which yields:

$$x_+ = x_c - (J(x_c)^T J(x_c) + S(x_c))^{-1} J(x_c)^T R(x_c) \tag{3.28}$$

The iterative method which is based on this equation is usually called the "full Newton method"; it is a fast local method for the nonlinear least-squares problem [DS83]. However, due to the fact that $S(x)$ is usually either unavailable or inconvenient to obtain analytically, this method is not commonly used in practice.

To come up with a practical solution to the nonlinear least-squares problem, two alternatives are available: The first is based on approximating $R(x)$ using equation (3.24) and then minimizing

$$\hat{m}_c(x) = \frac{1}{2}\|M_c(x)\|_2^2$$

and the second is based on using equation (3.28) along with approximated $S(x_c)$. The first alternative leads to a linear least-squares problem which can be solved using the following equation to obtain the next iterate x_+:

$$x_+ = x_c - (J(x_c)^T J(x_c))^{-1} J(x_c)^T R(x_c) \tag{3.29}$$

The iterative method that consists of using equation (3.29) at each iteration is called the Gauss-Newton method. Equations (3.29) and (3.28) differ only in the term $S(x)$, included by Newton's method in the second derivative matrix given by equation (3.27) but omitted by the Gauss-Newton method. Equivalently, the only difference between the quadratic model $\hat{m}_c(x_c)$ and $m_c(x_c)$ given by equations (3.27) is that the portion $S(x_i)$ of $\nabla^2 f(x_c)$ is omitted from $\hat{m}_c(x_c)$. It can be shown [DS83] that if $S(x_*)$ is small relative to $J(x_*)^T J(x_*)$, where x_* is the solution, the Gauss-Newton method is locally or linearly convergent, i.e.,

$$|x_{k+1} - x_*| \leq c|x_k - x_*|$$

where $c \in [0, 1)$.

The main drawback of the Gauss-Newton method is that it may take steps that are too long, but in the correct direction. This suggests two ways of improving the Gauss-Newton algorithm: using it with a line search, or with a trust region strategy [DS83]. The first approach leads to

$$x_+ = x_c - \lambda_c (J(x_c)^T J(x_c))^{-1} J(x_c)^T R(x_c) \tag{3.30}$$

where λ_c is chosen in such a way that $f(x_+)$ decreases adequately. In the trust region approach, the following problem must be solved:

$$\begin{gathered} \text{Minimize } \|R(x_c) + J(x_c)(x_+ - x_c)\|_2^2 \\ \text{subject to } \|x_+ - x_c\| \leq \delta_c. \end{gathered}$$

This minimization problem is based on the idea of using an approximated linear model to $R(x)$ along with δ_c, an upper bound step length, for obtaining the best subsequent point to the solution. This leads to the following equation which forms the Levenburg-Marquardt method [DS83]:

$$x_+ = x_c - (J(x_c)^T J(x_c) + \mu_c I)^{-1} J(x_c)^T R(x_c) \tag{3.31}$$

where

$$\begin{cases} \mu_c = 0 & \text{if } \|(J(x_c)^T J(x_c))^{-1} J(x_c)^T R(x_c)\|_2 \leq 0 \\ \mu_c > 0 & \text{otherwise} \end{cases}$$

The other alternative to find a practical method to solve the least-squares problem is based on approximating $S(x)$ in equation (3.37) rather than completely omitting it; this leads to the full-Newton type methods which are slow, though they can be superior to the Gauss-Newton method on large residual problems. One way to obtain approximation is using a second approach that approximates $S(x_c)$ by $A(x_c)$ which is updated iteratively based on the following equation:

$$A_+(x_+ - x_c) = J(x_+)^T R(x_+) - J(x_c)^T R(x_+) \tag{3.32}$$

3.6 Large Scale Sparse Nonlinear Least-Squares Problems

In some applications of seismic tomography [BBC+85], a kind of nonlinear least-squares problems arises where the Jacobian matrix of the system is large and sparse. To solve this kind of problems, Martinez [Mar87] introduced an algorithm in which the Gauss-Newton method is approximately solved at each iteration using an iterative preconditioned conjugate gradient algorithm, then a new point is obtained using a two-dimensional trust region scheme. Reported results indicate that this algorithm is promising for large-scale problems. The following is the step by step algorithm as described in [Mar87].

To minimize $\frac{1}{2}\|R(x)\|_2^2$, let x^0 be an arbitrary initial point, (η_k) a sequence of strictly positive numbers such that limit $\eta_k = 0$, $\Theta_1, \Theta_2 \in (0,1)$, $\Theta_3 \in (0, \frac{1}{2})$, and $0 \leq \underline{M} < \overline{M} < \infty$.

Let x^k be the kth approximation to the solution obtained by the algorithm. We denote

$$R_k = R(x^k),\ J_k = J(x^k),\ g_k = J_k^T R_k = \left. \nabla(\tfrac{1}{2}\|R(x)\|^2) \right|_{x^k}.$$

To obtain x^{k+1}, in the case $g_k \neq 0$, we perform the following steps:

Step 1: compute J_k and g_k. If $g_k = 0$, stop.

Step 2: Obtain $w_k \in \mathbf{R}^n$ such that $\|J_k^T J_k w_k + g_k\| \leq \eta_k \|g_k\|$

Step 3: Obtain $v_k \in \mathbf{R}^n$ as the solution of the following bidimensional problem:

$$\begin{aligned} &\text{Minimize} \quad \|J_k v + R_k\| \\ &\text{subject to} \quad v = \lambda_1 g_k + \lambda_2 w_k, \|v\| \leq \|w_k\| \end{aligned}$$

Step 4: Set $d_k^1 = -g_k$. test the following conditions for v_k:

$$< v_k, g_k > \leq -\Theta_1 \|v_k\| \|g_k\|$$

and

$$\underline{M}\|g_k\| \leq \|v_k\| \leq \overline{M}\|g_k\|$$

If these inequalities are satisfied, set $d_k^2 = v_k$, otherwise $d_k^2 = d_k^1$.

Step 5: Set $t = \|d_k^2\|$. Perform the following steps:

(a) obtain d as the solution of the problem:

$$\text{Minimize } \|J_k d + R_k\|$$
$$\text{subject to } \; d = \lambda_1 d_k^1 + \lambda_2 d_k^2, \|d\| \leq t$$

(b) If

$$\frac{1}{2}\|R(x^k + d)\|^2 \leq \frac{1}{2}\|f(x^k)\|^2 + \Theta_2 < g_k, d > ,$$

go to (d).

(c) Let $\bar{t}$ be such that $\Theta_3 t \leq \bar{t} \leq (1 - \Theta_3)t$, replace t by $\bar{t}$ and go to step (a).

(d) $d_k = d$, $x^{k+1} = x^k + d_k$.

3.7 Constrained Optimization

The general nonlinear programming problem may be stated in the following form:

$$\text{NP}: \; \text{minimize} f(x) \; \text{ subject to } l \leq \begin{Bmatrix} x \\ A_L x \\ C(x) \end{Bmatrix} \leq u$$
$$x \in \mathbf{R}^n$$

where $f(x)$ is a smooth nonlinear objective function, A_L is a constant matrix of constraints, and $C(x)$ is a vector of smooth nonlinear constraint functions. It is widely agreed that problems with nonlinear constraints are considerably more difficult than those with purely linear constraints. In general, the difficulty arises because of the conflict between unconstrained reducing f and satisfying the constraints. Methods for solving problems were briefly described in previous sections; on the other hand, variable bound and/or general linear constraints add few conceptual difficulties to the basic iterative procedures of unconstrained problems [GMS+88]. In contrast, attaining and maintaining feasibility with respect to nonlinear constraints are potentially very difficult tasks. There is no guaranteed procedure to determine whether a feasible point exists with respect to a set of general nonlinear constraints; furthermore, even if a feasible point is given, moving to another feasible point generally requires an iterative procedure. Thus, any method for NP must somehow balance changes in f against changes in the constraint violations.

The general approach in nonlinear optimization methods is to transform NP into a sequence of solvable sub-problems. Due to the relative ease with which unconstrained problems can be solved, most methods developed earlier attempted to transform NP into an unconstrained sub-problem. We next give a brief description of some of the well-known methods of nonlinear constraint optimization. We focus on one method that is widely regarded as the most effective general method for NP. Since procedures for variable bound and linear constraints are relatively well understood, we consider only nonlinear constraints, i.e., NP is defined as:

$$\begin{aligned} \text{NP}: \ & \text{minimize} f(x) \ \text{ subject to } C(x) \geq 0 \\ & x \in \mathbf{R}^n \end{aligned} \tag{3.33}$$

3.7.1 Penalty-Function Methods

A penalty function is a combination of the original objective function and a penalty term that increases monotonically with some measure of constraint violation. The most widely used differentiable penalty term is the sum of squared constraint violations, which gives the quadratic penalty function:

$$P_q(x,\rho) = f(x) + \frac{1}{2}\rho\sum_{i=1}^{m}[min(0,c_i(x))]^2 \tag{3.34}$$

Let $x^*(\rho)$ denote the unconstrained minimum of $P_q(x,\rho)$; under mild conditions, it can be shown that $x^*(\rho)$ approaches x^* as ρ approaches infinity. Thus, the classical quadratic penalty function approach consists of finding the unconstrained minimum of $P_q(x,\rho)$ for a sequence of increasing values of ρ. Methods based on this approach are often called "sequential unconstrained minimization techniques". Unfortunately, quadratic penalty function methods suffer from certain difficulties. In particular, the penalty parameter ρ in equation (3.34) must approach infinity for $x^*(\rho)$ to converge to x^*. Each increase in ρ leads to increased ill-conditioning in the unconstrained sub-problems. Furthermore, these methods are inefficient, if each sub-problem is solved accurately [GMS+88].

These difficulties can be avoided by using a non-differentiable penalty function, of which the most popular is the l_1 penalty function:

$$P_1(x,\rho) = f(x) + \rho\sum_{i=1}^{m}|min(0,c_i(x))| \tag{3.35}$$

Under mild conditions, there is a finite threshold value ρ such that x^* is a local minimum of P_1 for $\rho > \overline{\rho}$ [GMS+88]. This implies that x^* can be computed with a single unconstrained minimization; for this reason, penalty functions like P_1 are sometimes termed exact penalty functions.

3.7.2 Barrier Function Method

An effective approach to deal with inequality constraints is based on the use of barrier functions, which combines the original objective function with a weighted barrier term, a function with a positive singularity at the boundary of the feasible region. The most popular are the logarithmic and inverse barrier functions:

$$B_L(x,\mu) = f(x) - \mu \sum_{i=1}^{m} L_n c_i(x)$$

and

$$B_I(x,\mu) = f(x) + \mu \sum_{i=1}^{m} \frac{1}{c_i(x)}$$

Let $x^*(\mu)$ denote the unconstrained minimum of $B(x,\mu)$; under mild conditions, it can be shown that $x^*(\mu)$ approaches x^* as μ approaches zero [GMS+88]. If a strictly feasible starting point is known, a barrier function method will produce a sequence of improving (and strictly interior) estimates of x^*. However, the extreme nonlinearity and special properties of barrier functions make it difficult to apply standard unconstrained methods.

3.7.3 Generalized Reduced-Gradient Methods

Another approach to treating nonlinear constraints is based on the idea of reducing f while remaining on the nonlinear constraints. The best known such methods are called "generalized reduced-gradient (GRG) methods". The sub- problem in a GRG method usually involves changing x in order to reduce f, followed by an iterative procedure to restore feasibility.

GRG methods have been effective on many practical problems. However, they can be extremely inefficient when applied to problems with highly nonlinear con-

straints; they have tended to be most successful when applied to constraints that are almost linear.

3.7.4 Sequential Quadratic Programming Methods

Sequential quadratic programming (SQR) methods are widely regarded as some of the most efficient methods for solving practical optimization problems. As presented by Schittkowski [Sch88], suppose that we want to solve a general nonlinear programming problem of the type

$$\begin{aligned} &\text{Minimize} \quad f(x) \\ &x \in \mathbf{R}^n \\ &\text{subject to} \begin{cases} g_j(x) = 0 & , j = 1, \cdots, m_e \\ g_j(x) \geq 0 & , j = m_e + 1, \cdots, m \end{cases} \end{aligned} \tag{3.36}$$

with continuously differentiable real-valued functions f and $g_1, \cdots, g_m$.

The basic idea is to formulate and solve, in each iteration, a quadratic programming sub-problem, which is obtained by linearizing the constraints and approximating the Lagrange function

$$L(x, u) = f(x) - \sum_{j=1}^{m} u_j g_j(x) \tag{3.37}$$

quadratically, where $x \in \mathbf{R}^n$ and $u = (u_1, \cdots, u_m)^T \in \mathbf{R}^m$ is the vector of Lagrange multipliers.

Let $x_k \in \mathbf{R}^n$, $v_k \in \mathbf{R}^m$ and $B_k \in \mathbf{R}^{n \times n}$ denote the approximated solution, Lagrange multipliers and Hessian of Lagrange function, respectively. Thus, we

must solve the following quadratic programming problem:

$$\begin{gathered} \text{Minimize } \tfrac{1}{2}d^T B_k d + \nabla f(x_k)^T d \\ x \in \mathbf{R}^n \\ \text{subject to} \begin{cases} \nabla g_j(x_k)^T d + g_j(x_k) = 0 & ,j = 1, \cdots, m_c \\ \nabla g_j(x_k)^T d + g_j(x_k) \geq 0 & ,j = m_c + 1, \cdots, m \end{cases} \end{gathered} \tag{3.38}$$

Let d_k be the solution of this sub-problem and u_k the corresponding multiplier. A new iterate is obtained by

$$\begin{Bmatrix} x_{k+1} \\ v_{k+1} \end{Bmatrix} = \begin{Bmatrix} x_k \\ v_k \end{Bmatrix} + \alpha_k \begin{Bmatrix} d_k \\ u_k - v_k \end{Bmatrix} \tag{3.39}$$

where $\alpha_k \in (0, 1]$ is a suitable step-length parameter. Although it can be guaranteed that the matrix B_k is positive definite, it is possible that (3.38) is not solvable due to inconsistent constraints. One possible remedy is to introduce an additional variable $\delta \in \mathbf{R}$, leading to the modified problem:

$$\begin{gathered} \text{Minimize } \tfrac{1}{2}d^T B_k d + \nabla f(x_k)^T d + \rho_k \delta^2 \\ x \in \mathbf{R}^n \\ \text{subject to} \begin{cases} \nabla g_j(x_k)^T d + (1-\delta) g_j(x_k) \geq 0 & ,j \in J_k \\ \nabla g_j(x_k)^T d + g_j(x_k) \geq 0 & ,j \in K_k \end{cases} \\ 0 \leq \delta \leq 1 \end{gathered} \tag{3.40}$$

where J_k, the active set is given by:

$$J_k = \{1, \cdots, m_e\} \cup \{j | m_e < j \leq m,\ g_j(x_k) < \epsilon \text{ or } u_j^{(k)} > 0\}$$

and K_k is the complement set. In (3.40), ϵ is any small tolerance to define the active constraints and $u_j^{(k)}$ denotes the jth coefficient of u_k. Obviously, the point $(d_0 = 0, \delta = 1)$ satisfies the linear constraint of (3.40) which is therefore always

solvable. To avoid calculation of second derivatives and to obtain a final super-linear convergence rate, the standard approach is to update B_k as follows:

$$B_{k+1} = B_k + \frac{q_k q_k^T}{q_k^T p_k} - \frac{B_k p_k p_k^T B_k}{p_k^T B_k p_k} \tag{3.41}$$

where

$$q_k = \nabla_x L(x_{k+1}, u_k) - \nabla_x L(x_k, u_k)$$

and

$$p_k = x_{k+1} - x_k$$

3.8 Constrained Nonlinear Least-Squares Problem

This problem can be defined in the same way as (3.36) where

$$f(x) = \frac{1}{2} R(x)^T R(x).$$

Since most nonlinear least-squares problems are ill-conditioned, general nonlinear programming methods may not handle them properly. One alternative to solve this problem is to use an algorithm that is designed for this purpose. However, very few algorithms have been presented in the literature to tackle this problem. Mahdavi Amiri [MA81] proposed an approach based on an SQP method. Another algorithm was proposed by Mahdavi-Amiri and Bartles [MAB89] which is based on an exact penalty function. Schittkowski [Sch88] showed that a simple transformation of the original constrained nonlinear least-squares problem and its subsequent solution by a general purpose SQP algorithm retains typical features of a special purpose code and prevents the need to take care of any negative eigen values of an approximated

Hessian matrix. Due to the availability of many general purpose SQP codes and the simplicity of implementing the transformation, this idea is appealing. The transformation is performed by introducing l additional variables $y =< y_1, \cdots, y_l >^T$ and l additional equality constraints of the form:

$$f_i(x) - y_i = 0, i = 1, \cdots, l \tag{3.42}$$

Considering the unconstrained least-squares problem (3.23) first, the equivalent transformed problem is:

$$\begin{aligned} &\text{Minimize } \tfrac{1}{2} y^T y \\ &< x, y >^T \in \mathbf{R}^{n+l} \\ &\text{subject to } R(x) - y = 0 \end{aligned} \tag{3.43}$$

This can be written in a general form as follows:

$$\begin{aligned} &\text{Minimize } \bar{f}(\bar{x}) \\ &\bar{x} \in \mathbf{R}^{\bar{n}} \\ &\text{subject to } \bar{g}(\bar{x}) = 0 \end{aligned} \tag{3.44}$$

where $\bar{n} = n + l$, $\bar{x} =< x, y >^T$, $\bar{f}(\bar{x}) = \frac{1}{2} y^T y$ and $\bar{g}(\bar{x}) = R(x) - y$. Applying the SQP method derived in Section 3.7 leads to the following quadratic programming sub-problem:

$$\begin{aligned} &\text{Minimize } \tfrac{1}{2} \bar{d}^T \bar{B}_k \bar{d} + \nabla \bar{f}(\bar{x}_k)^T \bar{d} \\ &\bar{d} \in \mathbf{R}^{\bar{n}} \\ &\text{subject to } \nabla \bar{g}(\bar{x}_k)^T \bar{d} + \bar{g}(\bar{x}_k) = 0 \end{aligned} \tag{3.45}$$

where $\bar{B}_k$ is an approximation of the Hessian of the Lagrange function

$$L(\bar{x}, u) = \bar{f}(\bar{x}) = u^T \bar{g}(\bar{x})$$

which can be obtained as follows:

$$\bar{B}_k = \begin{bmatrix} B_k & 0 \\ 0 & I \end{bmatrix} \tag{3.46}$$

Substituting (3.46) into (3.45) gives the equivalent quadratic programming sub-problem:

$$\begin{aligned} &\text{Minimize } \tfrac{1}{2}d^T B_k d + \tfrac{1}{2}e^T e + y_k^T e \\ &< d, e >^T \in \mathbf{R}^{n+l} \\ &\text{subject to } \nabla R(x_k)^T d - e + R(x_k) - y_k = 0 \end{aligned} \tag{3.47}$$

If we take nonlinear constraints into account, the following transformed problem must be solved by an SQP method:

$$\begin{aligned} &\text{Minimize } \tfrac{1}{2}y^T y \\ &y \in \mathbf{R}^l \\ &\text{subject to} \begin{cases} f_i(x) - y_i = 0 & ,i = 1, \cdots, l \\ g_j(x) = 0 & ,j = 1, \cdots m_e \\ g_j(x) \geq 0 & ,j = m_e + 1, \cdots, m \\ x_l \leq x \leq x_u \end{cases} \end{aligned} \tag{3.48}$$

In this case the quadratic programming sub-problem is of the kind:

$$\begin{aligned} &\text{Minimize } \tfrac{1}{2}\begin{pmatrix} d^T & e^T \end{pmatrix} \bar{B}_k \begin{pmatrix} d \\ e \end{pmatrix} + y_k^T e \\ &< d, e >^T \in \mathbf{R}^{n+l} \\ &\text{subject to} \begin{cases} \nabla f_i(x_k)^T d - e_i + f_i(x_k) + y_k^{(k)} = 0 & ,i = 1, \cdots, l \\ \nabla g_j(x_k)^T d + g_j(x_k) = 0 & ,j = 1, \cdots m_e \\ \nabla g_j(x_k)^T d + g_j(x_k) \geq 0 & ,j = m_e + 1, \cdots, m \\ x_l - x_k \leq d \leq x_u - x_k \end{cases} \end{aligned} \tag{3.49}$$

This is basically identical to the minimization problem 3.38.

3.9 Sparse Least-Squares Problem

Linear least-squares problems arise frequently in many applications in tomography where data fitting usually forms the core of the inverse problem. In linear data fitting, the least-squares problem appears as the major part of the inverse process, while in nonlinear data fitting, nonlinear least-squares problem arises and must be solved iteratively; each iteration usually includes a least-squares problem. Thus, having an efficient algorithm for solving the linear least-squares problem is crucial for most inverse problems. In this Section we will describe four methods, some of them are widely used to solve the least-squares problem.

3.9.1 The Method of Normal Equations

The most widely used method for solving the full rank least-squares problem is the method of normal equations.

For $A^TAx = A^Tb$, where $A \in \mathbf{R}^{n\times n}$ with the property that $rank(A) = n$ and $b \in \mathbf{R}^n$ we can summarize the algorithm of normal equations to compute the least-squares solution as follows:

- Compute the lower triangular portion of $C = A^TA$.
- Compute the Cholesky factorization $C = GG^T$.
- Solve $Gy = d$. $(d = A^Tb)$
- Solve $G^Tx_{LS} = y$.

The above algorithm is convenient because it relies on standard algorithms: Cholesky factorization and matrix multiplication. The sensitivity of the x_{LS} solution obtained from this algorithm depends on $k_2(A)$ = square of the ratio of the largest

to the smallest singular values. Therefore, when k_2 is large, the normal equations method gives inaccurate results. In large sparse matrices, the first step of this method can be performed efficiently using the fast algorithms presented by Gustavson [Gus78] to obtain A^T and A^TA of a sparse matrix A. This algorithm can be applied for sparse matrices stored using a row-wise storing scheme usually called "IA-JA scheme". In this scheme a matrix A with m rows and n columns is stored using three one-dimensional arrays. Two arrays list the column indices (JA) and numerical values (A) of the NA non-zero matrix entries. The arrangement of these elements is row-wise. The third array is a set of row (address) pointers (IA) where ith element of IA is the address in both JA and A of the first non-zero element in the ith row of A.

3.9.2 QR Factorization

Another possible way to solve the least-squares problem is applying the QR factorization. For a given $A \in \mathbf{R}^{m\times n}$ with $m \geq n$ and $b \in \mathbf{R}^m$, suppose that an orthogonal matrix $Q \in \mathbf{R}^{m\times m}$ has been computed such that

$$Q^TA = R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}_{m-n}^{n}$$

is upper triangle. If

$$Q^Tb = \begin{bmatrix} c \\ d \end{bmatrix}_{m-n}^{n}$$

then, since Q is orthogonal, we have

$$\|Ax - b\|_2^2 = \|Q^TAx - Q^Tb\|_2^2 = \|R_1x - C\|_2^2 + \|d\|_2^2$$

for any $x \in \mathbf{R}^n$. If $rank(A) = rank(R_1) = n$, the least-squares solution is defined by the upper triangular system:

$$R_1 x_{LS} = c$$

Therefore, the full rank least-squares problem can be readily solved once the QR factorization of A is computed. There are many algorithms to carry out the QR factorization such as Householder QR factorization which utilizes Householder transformations [GVL89]. The Householder approach produces inaccurate results when applied to large residual, ill-conditioned problems; however, it is applicable to wider class of matrices compared to the method of normal equations because the Cholesky process applied to $A^T A$ breaks down before the back substitution process on $Q^T A = R$ [GVL89].

3.9.3 LSQR Algorithm

The LSQR algorithm by Paige and Saunders [PS82] has been applied successfully in various areas of inverse problem. The algorithm is basically designed to solve the least-squares problem especially when A, the coefficient matrix is large and sparse, and it is based on the bidiagonalization procedure of Golub and Kahan [GK65]. It generates a sequence of approximations x_k such that the residual norm $\|r_k\|_2$ decreases monotonically, where $r_k = b - Ax_k$. This algorithm is similar in style to the well-known Conjugate Gradient method and can be summarized as follows:

1. Initialize $\beta_1 u_1 = b$, $\alpha_1 v_1 = A^T u_1$, $w_1 = v_1$, $x_0 = 0$, $\overline{\Phi} = \beta_1$, and $\overline{\rho} = \alpha_1$.

2. For $i = 1, 2, 3, \cdots$ repeat steps 3-6

3. Continue the bidiagonalization

(a) $\beta_{i+1}u_{i+1} = Av_i - \alpha_i u_i$

(b) $\alpha_{i+1}v_{i+1} = A^T u_{i+1} - \beta_{i+1}v_i$

4. Construct and apply next orthogonal transformation:

(a) $\rho_i = (\overline{\rho}_i^2 + \beta_{i+1}^2)^{\frac{1}{2}}$

(b) $c_i = \frac{\overline{\rho}_i}{\rho_i}$

(c) $s_i = \frac{\beta_{i+1}}{\rho_i}$

(d) $\Theta_{i+1} = s_i\alpha_{+1}$

(e) $\overline{\rho}_{i+1} = -c_i\alpha_{i+1}$

(f) $\Phi_i = c_i\overline{\Phi}_i$

(g) $\overline{\Phi}_{i+1} = s_i\overline{\Phi}_i$

5. Update x and w

(a) $x_i = x_{i-1} + \frac{\Phi_i}{\rho_i}w_i$

(b) $w_{i+1} = v_{i+1} - \frac{\Theta_{i+1}}{\rho_i}w_i$

6. Test for convergence. Exit if some stopping criteria have been met.

From numerical tests conducted by Paige and Saunders [PS82], it can be concluded that the above algorithm is a very good alternative to solve least- squares problems that arise in inverse problem applications because it showed a good performance when A is ill-conditioned.

3.9.4 Solving The Augmented System

The regularization technique is used broadly in various inverse problem applications; this technique, as shown earlier, leads to equation (3.12):

$$(G^T G + \mu^2 C^T C)x = G^T d.$$

In large scale inverse problems, both matrices G and C are large and sparse, and it is clear that

$$A = G^T G + \mu^2 C^T C$$

is symmetric, positive definite; however, the matrix-matrix multiplication involved to obtain A may lead to a dense matrix. To avoid the matrix-matrix multiplication, keep the sparsity of the original matrices and to save the required computation time considerably, it may be desired to solve the following augmented system:

$$\begin{bmatrix} I & G \\ G^T & -\lambda^2 C^T C \end{bmatrix} \begin{Bmatrix} r \\ x \end{Bmatrix} = \begin{Bmatrix} b \\ 0 \end{Bmatrix} \tag{3.50}$$

where $r = b - Gx$ is the residual. It can be shown that for commonly used C matrices, $C^T C$ is a sparse matrix, especially for large-scale problems. Thus, the matrix coefficient of equation (3.50) is still sparse; also it is symmetric, but indefinite. This kind of sparse systems can be solved very efficiently using sparse solvers, either directly [GL81] or iteratively [MV77].

3.10 Closure

In this chapter, we have presented a literature review of inverse problem methods in computerized tomography. In this section, we will summarize each presented method.

Methods of inversion based on the wave approach are reviewed in section 3.2. These methods are based on the governing wave p.d.e's that ultimately lead to nonlinear optimization problems. Because of the high nonlinearity of the governing equations, and on the other hand, the large number of unknowns involved, numerical optimization algorithms are cumbersome and often converge to a local minimum. For stress induced anisotropic media, the wave p.d.e's are not available, therefore this approach is not applicable in this study.

In section 3.3, we presented transform based inversion methods. These methods are based on either the Fourier diffraction theorem or the Fourier slice theorem. These method require illuminating the object by plane waves from multiple directions; this is very difficult in the borehole to borehole technique used in this study. Another difficulty in these methods is the extraction of the scattered wave field from the total recorded wave field. Finally, for stress induced anisotropic media, neither the Fourier diffraction theorem nor the Fourier slice theorem is applicable.

In section 3.4, we reviewed methods of ray based travel-time tomography that, due to their relative simplicity, are very attractive. However, all of those methods are only applicable for inhomogeneous media which are isotropic; for those media the line integral travel-time equation has linear kernel.

Ray based travel-time tomography methods in inhomogeneous anisotropic media are reviewed in section 3.4.3. Those methods are presented for inhomogeneous cross-isotropic media; they lead to iterative inversion schemes. For stress induced anisotropic media, no inversion method has been suggested up until now.

As we will find in the next chapter, the inversion models presented for stress induced anisotropic media lead to nonlinear least squares problems. In section 3.5, we reviewed the fundamentals of the unconstrained nonlinear least squares problem

in general, while in section 3.6, we addressed the same problem when large sparse Jacobian matrices are involved. The latter arise in the first inversion approach presented in the next chapter.

In section 3.7, we reviewed the basics of constrained nonlinear optimization problem in general. Among methods used to solve this problem, SQR methods are widely regarded as some of the most efficient. Based on these methods, we presented an efficient algorithm for solving constrained nonlinear least squares problems which arise in the second inversion approach presented in the next chapter.

Finally, in section 3.9, we reviewed some numerical issues that arise in sparse least squares problems.

Chapter 4

Mathematical Modeling

4.1 Introduction

In this chapter, we formulate a stress reconstruction technique of a 2-D granular soil medium by applying travel-time tomography. We assume that before it is loaded, the soil is homogeneous, or at least piece-wise homogeneous, and isotropic. Stress distribution corresponding to loading makes the medium inhomogeneous and anisotropic. In this formulation, we will adopt the elliptical anisotropy assumption. Treating the anisotropy in this way has the merit of providing a practical travel-time algorithm for highly anisotropic and inhomogeneous media.

The stress reconstruction technique of this study leads to an inverse problem. As in other inverse problems, the direct problem is an essential part of solving the corresponding inverse problem. Thus finding an efficient way to treat the direct problem is very important to achieve a practical method to solve the inverse problem. We will therefore adopt the ray approach to formulate the direct problem. We consider ray bending, which is very important in highly inhomogeneous media,

and the direct problem will be a matter of ray tracing and simple line integration along rays. For the ray tracing, we will present two methods. The first method basically extends an inhomogeneous-anisotropic medium of the method proposed by Asakawa and Kawanaka [AK93]. The second method is a bending method that uses the simplex method of Nelder and Mead [NM65] to find optimum ray paths.

For the inverse problem, we will present two methods based on two different approaches. In the first approach, which we call the *nodal stress approach*, we define the stress field in terms of nodal stress components. With this approach, we obtain a set of nonlinear equations that relates measured travel-times to the stress field. Since these equations are usually inadequate for determining the unknown stress field, we therefore impose the equilibrium equations to decrease the ill-conditioning of the equations. Furthermore, to obtain a completely stable system, we use a regularization technique. In this way, we obtain a large-scale nonlinear system that can be solved using an optimization technique.

In the second approach to the inverse problem, which we call the *stress potential function approach*, we adopt the idea of stress potential functions to represent the stress field. This approach has two advantages: First, it requires a small number of unknowns to define a stress field with a reasonable accuracy, and second, it defines stress components in such a way that they automatically satisfy equilibrium conditions. With this approach, we obtain a system of highly nonlinear equations, but with only a few unknowns. The system can be solved using optimization techniques; however, because of the high nonlinearity, it is quite possible that optimization leads to an undesired local minimum. To avoid this problem, we impose some constraints such as the boundary conditions and the positiveness of the minimum principal stress. We thus obtain a constrained nonlinear least-squares problem that we solve using an appropriate optimization technique.

The last part of this chapter is devoted to the extension of the stress potential function approach, previously described for the 2-D case, to the 3-D case.

4.2 Ray Based Travel-Time Equation for Stress Induced Anisotropic-Inhomogeneous Media

Due to the relative simplicity and the attractive features of the ray-based approach discussed in Chapter 2, we choose this approach to formulate the problem. The fundamental equation giving the travel-time associated with a ray path R in a general media is

$$t = \int_R S dr \tag{4.1}$$

where $S = \frac{1}{V}$ is the ray, or group, slowness at each location on the ray path, and dr is the incremental length along the ray path. In a heterogeneous anisotropic medium, S is a function of r, the position vector, and Φ, the angle which represents the slowness direction. Thus,

$$t = \int_R S(r, \Phi) dr \tag{4.2}$$

To derive the travel-time equation on the basis of equation (4.2), we assume elliptical anisotropy for the slowness of P-wave propagation as shown in Figure 4.1.

In Figure 4.1, $S_1(r)$ and $S_2(r)$ represent the maximum and minimum slownesses at the point r; they will be called principal slownesses. The directions corresponding to the principal slownesses are assumed to coincide with the principal axes of stress. The latter assumption is consistent with the power relation

$$V_P = \alpha \sigma^{\beta}$$

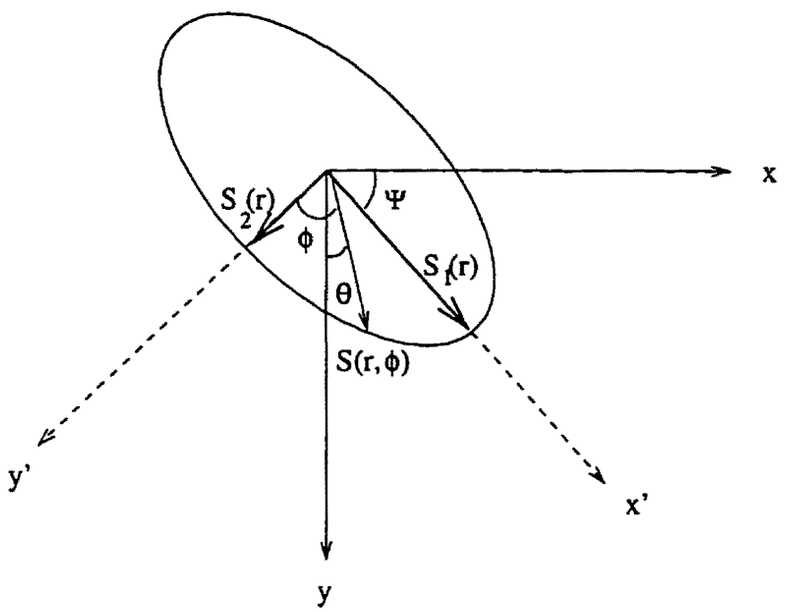


Figure 4.1: Elliptical anisotropy of slowness

described in Chapter 2. The following equation explains the elliptical directional variations of slowness:

$$S^2(r,\Phi) = S_1(r)^2 \sin^2\Phi + S_2(r)^2 \cos^2\Phi \tag{4.3}$$

Substituting $S(r,\Phi)$ from (4.3) in (4.2), we obtain the following equation in $x'y'$ rotating system:

$$t = \int_R \sqrt{S_1(r)^2(dx')^2 + S_2(r)^2(dy')^2} \tag{4.4}$$

Considering Ψ, the angle between the rotating $x'y'$ system and xy system, we have:

$$\begin{Bmatrix} dx' \\ dy' \end{Bmatrix} = \begin{bmatrix} \cos\Psi & \sin\Psi \\ -\sin\Psi & \cos\Psi \end{bmatrix} \begin{Bmatrix} dx \\ dy \end{Bmatrix} \tag{4.5}$$

Substituting dx' and dy' from (4.5) in (4.4), we obtain:

$$t = \int_R \sqrt{(\cos\Psi dx + \sin\Psi dy)^2 S_1(r)^2 + (-\sin\Psi dx + \cos\Psi dy)^2 S_2(r)^2} \tag{4.6}$$

Using equation (2.11) along principal stress directions leads to:

$$S_1(r) = \frac{1}{\alpha}\sigma_1^{-\beta} \tag{4.7}$$

$$S_2(r) = \frac{1}{\alpha}\sigma_2^{-\beta} \tag{4.8}$$

where σ_1 and σ_2 are the maximum and minimum principal stress and α and β are constants. Substituting $S_1(r)$ and $S_2(r)$ from these equations in (4.6) leads to the following fundamental travel-time equation which relates travel-time to stress field:

$$t = \frac{1}{\alpha}\int_R \sqrt{(\cos\Psi dx + \sin\Psi dy)^2\sigma_1^{-2\beta} + (-\sin\Psi dx + \cos\Psi dy)^2\sigma_2^{-2\beta}} \tag{4.9}$$

4.3 Ray Tracing in Stress Induced Inhomogeneous Anisotropic Media

In Chapter 2, we have briefly reviewed some ray tracing methods suggested for inhomogeneous-anisotropic media. All of the reviewed methods are based on the elastic wave equation in which the tensor of elastic parameters is used. Since at any point this tensor represents the inherent anisotropy of the medium of interest, these methods can therefore be used only in inherently anisotropic media; we are unaware of any ray tracing method for stress induced media.

In this section, based on the travel-time equation (4.9), we will present two methods for ray tracing in stress induced inhomogeneous anisotropic media.

4.3.1 A Forward-Backward Type Ray Tracing Method

This method essentially extends the linear travel-time interpolation method of ray tracing suggested by Asakawa and Kawanaka [AK93] for inhomogeneous isotropic

media. In this method, we divide the 2-D medium into a number of cells where we consider a number of extra points on the boundaries of each cell as depicted in Figure 4.2.

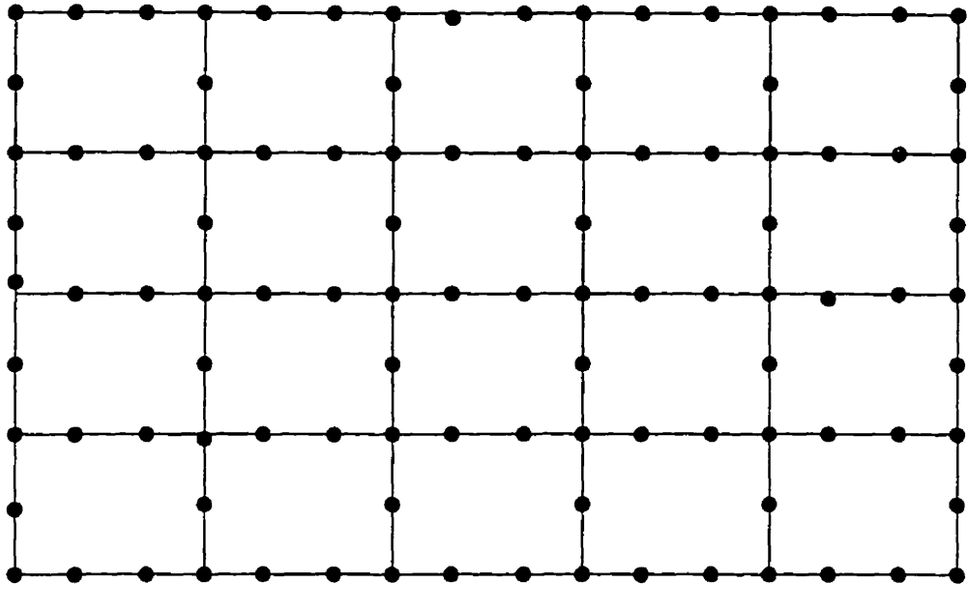


Figure 4.2: A typical grid used in ray tracing

The algorithm consists of two steps which we call *forward* and *backward* processes. In the forward process, for each source point we calculate travel-times at the discrete points. In the backward process, we trace ray paths between each pair of source and receiver points.

Basic Equations

To derive the basic equations, we consider a straight ray path segment AB and reaching point D as shown in Figure 4.3.

We seek a relation among t_A, t_B and t_D which are the travel-times at points A and B that are assumed to be known or computed in previous steps, and at

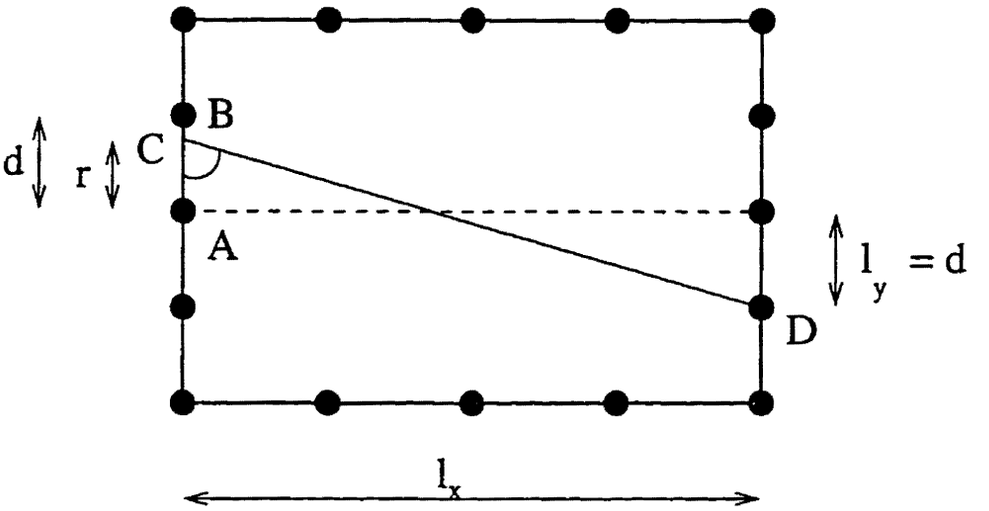


Figure 4.3: Segment AB of a ray path in a typical cell

an arbitrary discrete point D. We also need another relation to find the point C leading to minimum t_D. In terms of t_C, the travel-time t_D is given by

$$t_D = t_C + S(\theta)\sqrt{l_x^2 + (l_y + r)^2} \tag{4.10}$$

where l_x, l_y and r are shown in Figure 4.3, and $S(\theta)$ is the slowness along CD. Obtaining t_C by linear interpolation between t_A and t_B and using equations (4.3), (4.7) and (4.8), we obtain the following equation for t_D:

$$t_D = t_A\frac{d-r}{d} + t_B\frac{r}{d} + \frac{1}{\alpha}\sqrt{[\sigma_1^{-2\beta}\sin^2(\psi+\bar{\psi}) + \sigma_2^{-2\beta}\cos^2(\psi+\bar{\psi})][l_x^2 + (l_y+r)^2]} \tag{4.11}$$

where

$$\bar{\psi} = tan^{-1}(\frac{l_x}{l_y + r})$$

According to Fermat's principle, we obtain the correct travel-time t_D and the true ray by equating $\frac{\partial t_D}{\partial r}$ to zero. This leads to

$$\frac{l_x(\sigma_2^{-2\beta} - \sigma_1^{2\beta})}{l_x^2 + (l_y+r)^2}[\sin(\psi+\bar{\psi})\cos(\psi+\bar{\psi})][l_x^2 + (l_y+r)^2] +$$

$$(l_y + r)[\sigma_1^{-2\beta}\sin^2(\psi+\bar{\psi}) + \sigma_2^{-2\beta}\cos^2(\psi+\bar{\psi})] - \alpha\frac{(t_A - t_B)}{d} \times$$
$$\sqrt{[\sigma_1^{-2\beta}\sin^2(\psi+\bar{\psi}) + \sigma_2^{-2\beta}\cos^2(\psi+\bar{\psi})][l_x^2 + (l_y + r)^2]} = 0 \qquad (4.12)$$

We can obtain r by solving the above equation numerically, and then we can substitute it in equation (4.11) to obtain t_D.

Forward Process

In this process, we must calculate the travel-times corresponding to each source point at the discrete mesh points. To calculate the travel-time at a certain point, we must examine different possible segments, as AB in Figure 4.3, that the ray path may cross to reach the point. We then calculate the travel-times corresponding to each of these segments using equations (4.11) and (4.12) and choose the minimum one following the same procedure used by Asakawa and Kawanaka [AK93].

Backward Process

After we compute the travel-times on all cell boundaries that correspond to a certain shot point, we can find the ray paths by applying Fermat's principle. For each ray path we start from a receiver point; determine the position of point C using different possible segments on the boundary of the receiver cell; choose the one that corresponds to the minimum travel-time according to Fermat's principle. Having found C, we proceed by treating the resulting C as a new receiver point, until we reach the shot point. This procedure must be repeated for each receiver point to find the other ray paths.

4.3.2 Fermat's Principle-Based Ray Tracing Method

This method is essentially an extension to the bending type methods used by Prothero et al [PTE88] and Berryman [Ber90] for ray tracing in inhomogeneous isotropic media. To find a ray path between two points using this method, following Berryman [Ber90], we perturb the straight line connecting the two points by a number of algebraic terms such as sinusoidal or polynomial terms. These terms are multiplied by unknown coefficients that we find using Fermat's principle. According to "Weiestrass Polynomial Approximation Theorem" [LVG95], any function under mild conditions may be uniformly approximated arbitrarily closely by a polynomial. We therefore used a polynomial perturbation terms to find ray paths. Starting from the following polynomial of order n as a ray path

$$R(c) = \sum_{i=0}^{n} a_i x^i \tag{4.13}$$

for a ray path between a source and a receiver located at (x_1, y_1) and (x_2, y_2), by satisfying the end-points, it can be shown that the corresponding ray path can be expressed as

$$R(x) = y_1 + (x - x_1)\frac{y_2 - y_1}{x_2 - x_1} + \sum_{i=2}^{n} a_i[(x^i - x_1^i) - (x - x_1)\frac{x_2^i - x_1^i}{x_2 - x_1}] \tag{4.14}$$

From equation (4.9), we have

$$t = \frac{1}{\alpha}\int_R \sqrt{(\cos\Psi + \frac{dy}{dx}\sin\Psi)^2\sigma_1(r)^{-2\beta} + (-\sin\Psi + \frac{dy}{dx}\cos\Psi)^2\sigma_2(r)^{-2\beta}}dx \tag{4.15}$$

where $\frac{dy}{dx}$ of the ray path can be obtained from the following equation:

$$\frac{dy}{dx} = \frac{dR(x)}{dx} = \frac{y_2 - y_1}{x_2 - x_1} + \sum_{i=2}^{n} a_i[ix^{(i-1)} - \frac{x_2^i - x_1^i}{x_2 - x_1}] \tag{4.16}$$

According to Fermat's principle, energy travels along a ray which leads to the least travel-time; therefore considering a_i $(i = 2, \cdots, n)$ as variables, the ray tracing between (x_1, y_1) and (x_2, y_2) may be regarded as the following multivariable

optimization problem:

$$\text{Minimize} t = \frac{1}{\alpha}\int_R \sqrt{(\cos\Psi + \frac{dy}{dx}\sin\Psi)^2\sigma_1(r)^{-2\beta} + (-\sin\Psi + \frac{dy}{dx}\cos\Psi)^2\sigma_2(r)^{-2\beta}}dx \tag{4.17}$$

where

$$\frac{dy}{dx} = \frac{dR(x)}{dx} = \frac{y_2 - y_1}{x_2 - x_1} + \sum_{i=2}^{n} a_i[ix^i - \frac{x_2^i - x_1^i}{x_2 - x_1}] \tag{4.18}$$

The numerical implementation of this ray tracing method using the simplex method will be discussed in Chapter 5.

The two ray tracing methods presented in this chapter are both based on Fermat's principle, and therefore are capable of handling complicated wave phenomena such as diffraction. However, the first method is neither straightforward for programming nor is it efficient. The second method is on the other hand straightforward for numerical implementation and well-suited for inverse problems.

4.4 The Inverse Problem

The stress field of a soil medium can be represented in two ways: in the first, stress components are defined on a set of discrete pre-specified nodes as in the finite elements or finite difference methods of stress analysis; in the second, stress components of the entire medium ca be represented parametrically using suitable continuous functions. Based on these methods of stress representation, we will present two approaches to formulate the inverse problem

4.4.1 Nodal Stress Approach

In this approach, we define a set of nodes on the region of soil medium. Each node carries three unknowns that are the stress components required to define the state

of stress at a point in the 2-D case. The state of stress of any point in a 2-D problem can be represented either by:

1. the two principal stresses (σ_1, σ_2) and the angle they make with the x and y axes (Ψ) as shown in Figure 4.1; or
2. the three stress components in the Cartesian system (σ_x, σ_y, and τ).

The first way leads to highly nonlinear equilibrium equations. However, the second way has the advantage of leading to linear equilibrium equations as described shortly in this section.

In this approach, we use the travel-time equation in the Cartesian system to form a system of nonlinear equations that relates the unknown nodal stress to the measured travel-time. Because the number of unknown nodal stresses is usually larger than the number of measured travel-times, this system is inadequate for stress determination. We, therefore, add the equilibrium equations to the system. To stabilize, further, the system of travel-time and equilibrium equations, we use a regularization technique.

Travel-Time Equation in the Cartesian System

To represent stress in the Cartesian system, we use the stress sign convention illustrated in Figure 4.4 which is commonly used in soil mechanics problems.

The terms σ_1, $\sigma_2 \sin\Psi$ and $\cos\Psi$ which appear in equation (4.15) can be expressed through σ_x, σ_y and τ using the following known stress transformation equations:

$$\sigma_1 = \frac{\sigma_x + \sigma_y}{2} + \frac{1}{2}\sqrt{(\sigma_x - \sigma_y)^2 + 4\tau^2} \tag{4.19}$$

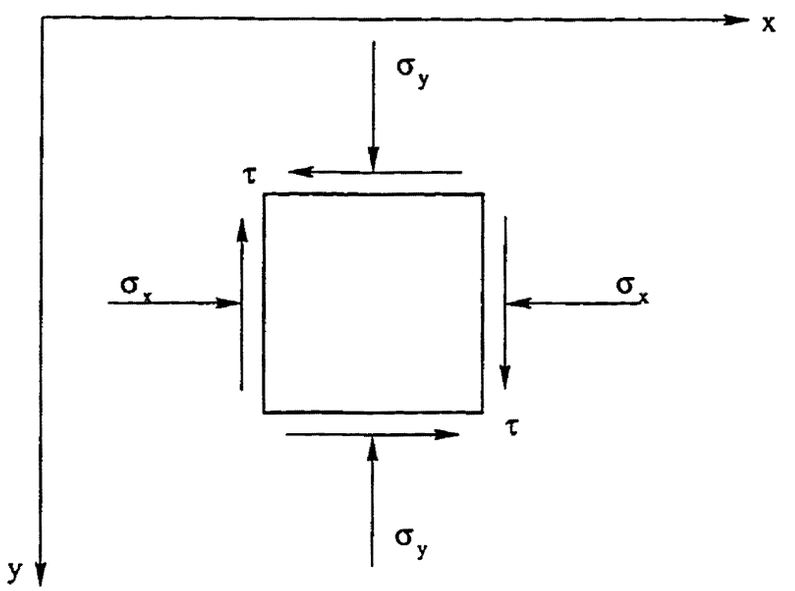


Figure 4.4: Stress Components

$$\sigma_2 = \frac{\sigma_x + \sigma_y}{2} - \frac{1}{2}\sqrt{(\sigma_x - \sigma_y)^2 + 4\tau^2} \tag{4.20}$$

$$\sin 2\Psi = \frac{2\tau}{\sqrt{(\sigma_x - \sigma_y)^2 + 4\tau^2}} \tag{4.21}$$

$$\cos 2\Psi = -\frac{\sigma_x + \sigma_y}{\sqrt{(\sigma_x - \sigma_y)^2 + 4\tau^2}} \tag{4.22}$$

Using equations (4.19) through (4.22), equation (4.15) leads to the following travel-time equation in terms of the stress components in the Cartesian system:

$$t = \frac{1}{\sqrt{2}\alpha}\int_R \sqrt{[1 + (\frac{dy}{dx})^2 - g]\sigma_1^{-2\beta} + [1 + (\frac{dy}{dx})^2 + g]\sigma_2^{-2\beta}}\,dx \tag{4.23}$$

where σ_1 and σ_2 are given in equations (4.19) and (4.20), and g is given by the equation:

$$g = \frac{(1 - (\frac{dy}{dx})^2)(\sigma_y - \sigma_x) + 4\frac{dy}{dx}\tau}{\sqrt{(\sigma_y - \sigma_x)^2 + 4\tau^2}} \tag{4.24}$$

Equation (4.23) relates each measured travel-time to stress components at points lying on the corresponding ray path. In Chapter 5 where the numerical imple-

mentation of the nodal stress approach is described, we will show how to obtain travel-time equations that relate measured travel-times to nodal stress components using an appropriate interpolation function.

Equilibrium Equations

Using the sign convention illustrated in Figure 4.4, we find that the force equilibrium requirements in the x and y directions lead to the following partial differential equations:

$$\frac{\partial \sigma_x}{\partial x} - \frac{\partial \tau}{\partial y} = 0 \tag{4.25}$$

$$\frac{\partial \tau}{\partial x} - \frac{\partial \sigma_y}{\partial y} = -\gamma_s \tag{4.26}$$

where γ_s is the density of the soil. In the nodal stress approach, these two equations lead to a set of linear equations relating nodal stress components of the soil medium this will be described in Chapter 5.

Regularization

With M measured travel-times and N nodes in a 2-D soil medium, the travel-time equation (4.23) provides M nonlinear equations, while the equations of equilibrium (4.25) and (4.26) each provides approximately N equations. Thus approximately $(2N + M)$ equations are available for obtaining $3N$ unknowns. The $2N$ equilibrium equations are independent while, due to the limited angles of shooting available in crosshole techniques, the M nonlinear travel-time equations are not completely independent. Therefore, the linear system which includes the equilibrium equations and any linearized version of the travel-time equations set, usually has a relatively high condition number.

Data noise is an invitable factor in this problem and we should therefore take proper measures to avoid undesired solutions. This factor leads to oscillatory solutions for the problem, because the travel-time equation is basically a Fredholm equation of the first type which is sensitive to data noise, especially when the linear system involved has a high condition number. The sensitivity to data noise in this problem is not expected to be as much as in common travel-time tomography, because the approximately $2N$ equilibrium equations are guaranteed to be independent.

In this approach, we use a regularization technique to avoid the undesired oscillatory solutions. This technique is based on adding some common sense constraints to the ill-conditioned system to secure its stability. For example, to avoid oscillatory solutions, a widely used method is to establish an expression that defines the function roughness, then to impose conditions for minimizing the expression. A typical expression for function roughness is Cx where C is usually a linear operator and x is the vector of unknowns. Thus, the regularization technique can be summarized as the additional information obtained from the following condition to avoid oscillatory solutions:

$$\text{Minimize} \quad \|C\sigma\| \tag{4.27}$$

where σ is the vector of unknown stresses in the Cartesian system.

Nonlinear Least Squares Solution

In the nodal stress approach, we obtain a set of nonlinear travel-time equations using equation (4.23) for each measured travel-time; this set can be represented as follows:

$$r_{n_i}(\sigma) = t_i - \frac{1}{\sqrt{2}\alpha}\int_{R_i}\sqrt{[(1+(\frac{dy}{dx})^2)-g]\sigma_1^{-2\beta}+[(1+(\frac{dy}{dx})^2)+g]\sigma_2^{-2\beta}}dx$$

$$\text{for } i = 1, \cdots, M \quad (4.28)$$

where $r_{n_i}(\sigma)$ denotes the nonlinear function of the ith travel-time residual. If we represent the linear equations obtained from equations (4.25) and (4.26) by

$$l_1(\sigma) = 0 \text{ and } l_2(\sigma) = 0$$

by introducing $r_{l_j}(\sigma)$ as the linear function of the jth equilibrium equation residual, we can obtain the following set of linear equations

$$r_{l_j}(\sigma) = l_j(\sigma), \quad j = M+1, \cdots, M+2N_c \quad (4.29)$$

where, if j is odd, $L_1(\sigma)$ must be used, otherwise $l_2(\sigma)$ must be used and N_c represents the number of cells used to discretize the medium. If we denote the residuals vector of equation (2.29) by $R_N(\sigma)$, and the residuals vector of equation (2.30) by $R_L(\sigma)$, considering the regularization strategy we can formulate the solution to the inverse problem as the solution of the following nonlinear least squares problem:

$$\text{Minimize } \frac{1}{2} \left\| \begin{array}{c} R_N(\sigma) \\ R_L(\sigma) \\ C\sigma \end{array} \right\|^2 \quad (4.30)$$

In Chapter 5, we will describe a numerical optimization technique for solving this nonlinear least square problem.

4.4.2 Stress Potential Function Approach

In this approach, instead of presenting the stress field by defining stress components on a set of discrete pre-specified nodes, we use appropriate continuous functions with only a few unknown parameters to represent the stress field of the entire

medium. Compared to the nodal stress approach, this approach has the advantage of envolving only few unknowns, so that it leads to a more reliable and less involved numerical inversion procedure. Another attractive feature of this approach is that we can define the stress functions in such a way that they automatically satisfy the equilibrium equations, therefore no additional equations of equilibrium are required.

The other attractive feature of this approach is the small sensitivity of the solution to data noise, because the number of measurements is much larger than the number of unknowns; regularization is usually not necessary. However, due to the high nonlinearity of the travel-time equations, it is likely that the numerical procedure of inversion in this approach will lead to an undesired solution. To avoid this problem, we impose some constraints as before.

Stress Potential Functions

Following the idea of Airy stress function [TG82], we can represent the three stress components in 2-D problems using only one potential function. In this way, in addition to considerably decreasing the number of unknown required parameters, the equations of equilibrium (4.25) and (4.26) are automatically satisfied. With Φ as a potential function, the three stress components can be represented in the following form:

$$\sigma_x = \frac{\partial^2 \Phi}{\partial y^2} + \gamma_s y \tag{4.31}$$

$$\sigma_y = \frac{\partial^2 \Phi}{\partial x^2} + \gamma_s y \tag{4.32}$$

$$\tau = \frac{\partial^2 \Phi}{\partial x \partial y} \tag{4.33}$$

where Φ can be any continuous function of x and y, however according to Weiestrass polynomial approximation theorem, polynomials are an appropriate choice for fit-

ting a stress field. Neglecting linear terms in a general two dimensional polynomial, we can use the following polynomial as a stress function:

$$\Phi(x,y,a) = a_1x^2 + a_2xy + a_3y^2 + a_4x^2 + a_5x^2y + a_6xy^2 + a_7y^3 + \cdots \quad (4.34)$$

where a represents the vector of unknown polynomial coefficients. Thus equations (4.31) to (4.32) show the stress components in terms of a are:

$$\sigma_x(x,y,a) = 2a_3 + 2a_6x + (6a_7 + \gamma_s)y + \cdots \quad (4.35)$$

$$\sigma_y(x,y,a) = 2a_1 + 6a_4x + (2a_5 + \gamma_s)y + \cdots \quad (4.36)$$

$$\tau(x,y,a) = a_2 + 2a_5x + 2a_6y + \cdots \quad (4.37)$$

Travel-Time Equation

The travel-time equation in this approach is exactly the same as equation (4.23) in which σ_1, σ_2 and g must be obtained by substituting σ_x, σ_y and τ from equations (4.35) to (4.37) in equations (4.19), (4.20) and (4.24), respectively.

Constraints

In this approach, we take advantage of some constraints to restrict the domain of the variables of the highly nonlinear governing equations. A number of constraints can be imposed due to the boundary conditions where some or all stress components are known. These constraints are linear in a and can be represented as follows:

$$C_{L_i}(a) = 0, \quad i = 1, \cdots, n_L \quad (4.38)$$

where n_L is the number of known stress components.

Another set of constraints can be obtained by considering the fact that non-cohesive granular soil media cannot resist tensile stress, therefore the minimum principal stress must be always positive. Thus from equation (4.20), we must have

$$\sigma_x + \sigma_y > \sqrt{(\sigma_x - \sigma_y)^2 + 4\tau^2} \tag{4.39}$$

where σ_x and σ_y must always be positive, otherwise this constraint cannot be satisfied, thus

$$\sigma_x > 0 \tag{4.40}$$

and

$$\sigma_y > 0 \tag{4.41}$$

On the other hand simplifying the inequality (4.39) leads to

$$\sigma_x \sigma_y - \tau^2 > 0 \tag{4.42}$$

Equations (4.40) and (4.41) lead to another set of linear constraints which can be represented as:

$$C_{L_j}(a) > 0, \quad j = 1, \cdots, 2n_p \tag{4.43}$$

where n_P is a number of points we can choose arbitrarily in the region. The inequality (4.42) leads to a set of nonlinear constraints which can be represented as follows:

$$C_{N_k}(a) > 0, \quad k = 1, \cdots, n_P \tag{4.44}$$

Constrained Nonlinear Least Squares Solution

As with the nodal stress approach, we can obtain a set of nonlinear travel-time equations represented by equation (4.28). Here, σ_1, σ_2 and g in the that equation can be written in terms of a the vector of unknown stress function coefficients; that

is, r_{n_i} is a function of a. Therefore, we denote the vector of residuals defined by equation (4.28) by $R(a)$.

Now, considering the linear equality constraints given by equation (4.38) and the linear and nonlinear inequality constraints given by (4.43) and (4.44), we can formulate the solution to the inverse problem as the solution to the following constrained nonlinear least squares problem:

$$\begin{aligned} &\text{Minimize } \tfrac{1}{2}\|R(a)\|^2 \\ &\text{subject to} \begin{cases} C_{L_i}(a) = 0 & ,i = 1,\cdots,n_L \\ C_{L_i}(a) > 0 & ,i = n_L+1,\cdots,n_L+2n_p \\ C_{N_i}(a) > 0 & ,i = n_L+2n_p+1,\cdots,n_L+3n_p \end{cases} \end{aligned} \tag{4.45}$$

In Chapter 5, we will describe a numerical optimization technique for solving this least squares problem.

4.5 A Method for the Velocity-Stress Power Relationship Parameters Determination

In Chapter 2, we presented equation (2.11) which is suggested by many researchers as the relation between P-wave velocity in particulate media, and the state of stress. For a specific soil media, the parameters α and β in that equation are determined experimentally. For this purpose, traditionally, a soil sample is provided and taken to a laboratory where it is subjected to a number of different isotropic stresses. Under each stress, the P-wave velocity is measured and recorded; with these measurements, α and β can be determined using either a linear regression analysis in log-log scale or a nonlinear least-squares algorithm; for more details see [Sch78, KSK82b, SP94, Cas96].

In this study, we suggest another method for the determination of α and β. This method is based on the in situ travel-time measurements; extraction of a soil sample, which leads to soil fabric alternation, is not required. Therefore, this method leads to more accurate parameters.

To determine α and β, in this method, we simply need to treat them as unknowns in the travel-time equation (4.2). In this way, we consider two more unknowns in the nonlinear leas-squares problem (4.30) or (4.45); this leads to the determination of α and β, as well as the stress field components.

4.6 The Stress Potential Function Approach in 3-D

In cases where loading and other circumstances may not justify a 2-D modeling, using a 3-D model is inevitable. To develop a 3-D model for soil stress reconstruction, one reasonable way is to use the same approach we used for the 2-D case while extending the methods and techniques involved in the forward modeling and the inverse problem where deemed necessary.

4.6.1 Travel-Time Equation in 3-D

As an extension to the ray based travel-time equation (4.9) obtained earlier in this chapter for the 2-D case, we will derive a similar travel-time equation that may be used in 3-D soil stress reconstruction model. Here, as an extension to the elliptical anisotropy used in the 2-D model, we assume an ellipsoidal anisotropy for the velocity of P-wave propagation. Thus, considering $V_1(r)$, $V_2(r)$ and $V_3(r)$ as

the principal velocities, in the direction of x', y' and z', a rotating system axes, we can use the following equation which describe the ellipsoidal directional variations of speed:

$$V(r,u)^{-2} = V_1(r)^{-2}e_1'^2 + V_2(r)^{-2}e_2'^2 + V_3(r)^{-2}e_3'^2 \tag{4.46}$$

Here $V(r,u)$ represents the P-wave speed at point r in the direction of the unit vector $u = u_1e_1' + u_2e_2' + u_3e_3'$, where (e_1', e_2', e_3') represents the vector basis in the $x'y'z'$ rotating system. In terms of slownesses, we obtain:

$$S(r,u)^2 = S_1(r)^2u_1'^2 + S_2(r)^2u_2'^2 + S_3(r)^2u_3'^2 \tag{4.47}$$

Considering dr' as the incremental vector along the ray path at point r, and its components in the $x'y'z'$ system, we have

$$dr' = dx'e_1' + dy'e_2' + dz'e_3' \tag{4.48}$$

Thus, we obtain the following equation by substituting (4.47) in the travel-time equation (4.2):

$$t = \int_R \sqrt{(S_1(r)^2u_1'^2 + S_2(r)^2u_2'^2 + S_3(r)^2u_3'^2)dr^2} \tag{4.49}$$

Since u and dr' are in the same direction, we have

$$\frac{dx'}{dr'} = u_1', \quad \frac{dy'}{dr'} = u_2', \quad \frac{dz'}{dr'} = u_3' \tag{4.50}$$

This leads to the following travel-time equation in the $x'y'z'$ system:

$$t = \int_R \sqrt{S_1(r)^2(dx')^2 + S_2(r)^2(dy')^2 + S_3(r)^2(dz')^2} \tag{4.51}$$

Using L, the well-known linear transformation matrix, we have:

$$\begin{Bmatrix} dx' \\ dy' \\ dz' \end{Bmatrix} = \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{33} \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{Bmatrix} dx \\ dy \\ dz \end{Bmatrix} \tag{4.52}$$

where l_{ij} $(i = 1, \cdots, 3,\ j = 1, \cdots, 3)$ are the components of L. Combining the last two equations leads to the following travel-time equation which can be used in ray tracing and the inverse problem:

$$\begin{aligned} t = \int_R [\ & (l_{11}dx + l_{12}dy + l_{13}dz)^2 S_1(r)^2 \\ & + (l_{21}dx + l_{22}dy + l_{23}dz)^2 S_2(r)^2 \\ & + (l_{31}dx + l_{32}dy + l_{33}dz)^2 S_3(r)^2\ ]^{\frac{1}{2}} \end{aligned} \tag{4.53}$$

or in terms of stress field using equations (4.7) and (4.8):

$$\begin{aligned} t = \frac{1}{\alpha} \int_R [\ & (l_{11}dx + l_{12}dy + l_{13}dz)^2 \sigma_1^{-2\beta} \\ & + (l_{21}dx + l_{22}dy + l_{23}dz)^2 \sigma_2^{-2\beta} \\ & + (l_{31}dx + l_{32}dy + l_{33}dz)^2 \sigma_3^{-2\beta}\ ]^{\frac{1}{2}} \end{aligned} \tag{4.54}$$

4.6.2 Ray Tracing in Stress Induced Inhomogeneous Anisotropic Media

In the 2-D case, we presented two methods for ray tracing, both of which are fundamentally extendible to the 3-D case. However, it is clear that the second method is more efficient, thus more suitable for the 3-D case where efficiency is yet more important than in the 2-D case. To extend that method to the 3-D case, we use parametric representation for ray paths. For a ray path between two points (x_s, y_s, z_s) and (x_r, y_r, z_r), we can use the following parametric equation:

$$\begin{cases} x(t) &= x_s + (x_r - x_s)t + \sum_{i=2}^{n_x} a_i t^i &, 0 \le t \le 1 \\ y(t) &= y_s + (y_r - y_s)t + \sum_{i=2}^{n_y} b_i t^i &, 0 \le t \le 1 \\ z(t) &= z_s + (z_r - z_s)t + \sum_{i=2}^{n_z} c_i t^i &, 0 \le t \le 1 \end{cases} \tag{4.55}$$

Here a_i, b_i and c_i are the perturbation coefficient sets, and n_x, n_y, and n_z are the number of perturbation terms in each set, respectively. The combination of

equations (4.54) and (4.55) leads to the travel-time equation in 3-D. For ray tracing, the unknown coefficient sets a_i, b_i and c_i must be found; this can be done, similar to the 2-D case, using the simplex method of Nelder and Mead [NM65].

4.6.3 The Inverse Problem

In the 3-D case, we can employ the idea of the stress potential function to formulate the inverse problem. In 3-D stress fields, there are six different stress components in the tensor of stress. Figure 4.5 shows stress components acting on a cubic element.

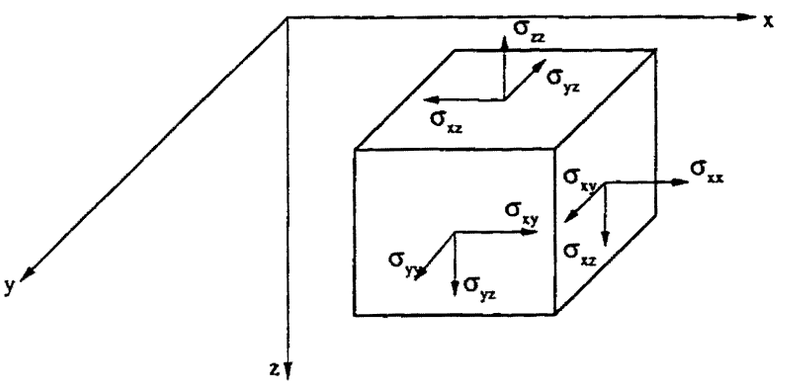


Figure 4.5: Stress Components

Considering gravity force in the z direction, the equilibrium equations can be obtained as follows:

$$\frac{\partial \sigma xx}{\partial x} + \frac{\partial \sigma xy}{\partial y} + \frac{\partial \sigma xz}{\partial z} = 0 \tag{4.56}$$

$$\frac{\partial \sigma xy}{\partial x} + \frac{\partial \sigma yy}{\partial y} + \frac{\partial \sigma yz}{\partial z} = 0 \tag{4.57}$$

$$\frac{\partial \sigma xz}{\partial x} + \frac{\partial \sigma yz}{\partial y} + \frac{\partial \sigma zz}{\partial z} + \gamma_s = 0 \tag{4.58}$$

Unlike in the 2-D case, unfortunately, the solution to these equations cannot be represented using a single potential function. However, to find stress functions that satisfy equilibrium, we can introduce a function with a number of unknown coefficients for each stress component. We then use equations (4.56) to (4.58) to find relations between the coefficients. For example, we can use the following polynomials to represent the stress components:

$$\begin{aligned}
\sigma_{xx} &= a_1 + a_2x + a_3y + a_4z + a_5x^2 + a_6y^2 + a_7z^2 + a_8xy + a_9xz + a_{10}yz + \cdots \\
\sigma_{yy} &= b_1 + b_2x + \cdots + b_{10}yz + \cdots \\
\sigma_{yy} &= c_1 + c_2x + \cdots + c_{10}yz + \cdots \\
\sigma_{xy} &= d_1 + d_2x + \cdots + d_{10}yz + \cdots \\
\sigma_{xz} &= e_1 + e_2x + \cdots + e_{10}yz + \cdots \\
\sigma_{yz} &= f_1 + f_2x + \cdots + f_{10}yz + \cdots
\end{aligned}$$

Substituting these functions in (4.56) through (4.58) leads to the following relations:

$$\begin{aligned}
a_2 &= -(d_3 + e_4) & b_3 &= -(d_2 + f_4) & c_4 &= -(e_2 + f_3 + \gamma_5) \\
a_5 &= -(d_8 + e_9)/2 & b_6 &= -(d_8 - f_{10})/2 & c_7 &= -(e_9 + f_{10})/2 \\
a_8 &= -(2d_6 + e_{10}) & b_8 &= -(2d_5 + f_9) & c_9 &= -(2e_5 + f_8) \\
a_9 &= -(d_{10} + 2e_7) & b_{10} &= -(d_9 + 2f_7) & c_{10} &= -(e_8 + 2f_6)
\end{aligned}$$

Therefore, in this example, we need $6 \times 10 - 3 \times 4 = 48$ independent unknowns to represent the stress field.

Chapter 5

Numerical Solution of the Model

5.1 Introduction

In Chapter 4, based on travel-time tomography, we presented a mathematical model for soil stress reconstruction. This model, basically, consists of two parts: the forward problem model, and the inverse problem model. The forward problem modeling lead to the fundamental travel-time equation (4.9). This equation is a line integral equation in which R, the integral path, is known. The main subject of the forward problem is the determination of the integration paths between any two points in a region with a given stress field. The process of determining these paths is called ray tracing. In this chapter, we will present a numerical technique for the ray tracing method based on Fermat's principle presented in Chapter 4.

The inverse problem is mainly governed by the same travel-time equation (4.9), but the subject here is the determination of stress field corresponding to a set of measured travel-times. In Chapter 4, we presented two approaches to formulate the inverse problem: the first led to a Large-scale nonlinear least-squares problem

(4.30) and the second led to a smaller constrained nonlinear least-squares problem. In this chapter, we will present detailed optimization techniques to solve these least-squares problems numerically.

5.2 Numerical Formulation

Here we present a numerical formulation for the ray tracing method based on Fermat's principle presented in Chapter 4.

In Chapter 4, we concluded that ray tracing between two points (x_1, y_1) and (x_2, y_2) in a region with given stress field leads to the following multivariable optimization problem (4.17):

$$\text{Minimize } t = \frac{1}{\alpha}\int_R \sqrt{(\cos\Psi + \frac{dy}{dx}\sin\Psi)^2\sigma_1^{-2\beta} + (-\sin\Psi + \frac{dy}{dx}\cos\Psi)^2\sigma_2^{-2\beta}}dr$$

where

$$\frac{dy}{dx} = \frac{y_2 - y_1}{x_2 - x_1} + \sum_{i=2}^{n} a_i(ix^{i-1} - \frac{x_2^i - x_1^i}{x_2 - x_1})$$

Solving this minimization problem leads to the determination of the unknown a_i's. To convert the integral equation to an algebraic equation, we introduce f_i as follows:

$$f_i = \sqrt{(\cos\Psi_i + (\frac{dy}{dx})_i\sin\Psi_i)^2\sigma_{1i}^{-2\beta} + (-\sin\Psi_i + (\frac{dy}{dx})_i\cos\Psi_i)^2\sigma_{2i}^{-2\beta}} \qquad (5.1)$$

Thus, using Simpson's rule, we obtain the following optimization problem:

$$\text{Minimize } t = \frac{\Delta x}{3\alpha}[f_1 + 4f_2 + 2f_3 + 4f_4 + \cdots + 2f_{N-2} + 4f_{N-1} + f_N] \qquad (5.2)$$

Here N is the number of discrete points we choose arbitrarily along the unknown ray path. The expression to be minimized in this minimization problem is a simple algebraic expression with a_i's, the unknown perturbation coefficients to be found. The

minimization problem (5.2) can be solved using any unconstrained multivariable minimization method. However, similar problems arise in ray tracing of other geographical applications are traditionally solved using the simplex method of Nelder and Mead [NM65]; see [PTE88, Ber90].

5.2.1 Ray Tracing Using the Simplex Method

The *simplex method* of multivariable minimization problems is due to Nelder and Mead [NM65]. The method requires only function evaluation and not derivatives. A simplex is the geometrical figure consisting, in n dimensions, of $n + 1$ points (or vertices) and all their interconnecting line segments, polygonal faces, etc. For example, in two dimensions a simplex is a triangle, and in three dimensions it is a tetrahedron.

The simplex method must be started with $n + 1$ points, defining an initial simplex. The initial simplex can be obtained by starting with an initial starting point P_0, then taking the other n points to be:

$$P_i = P_0 + \lambda e_i \tag{5.3}$$

where the e_is are n unit vectors, and λ is usually a constant that must be chosen properly.

Having an initial simplex, the simplex method takes a series of steps, most of that just move the point of the simplex where the function is largest (high point) through the opposite face of the simplex to a lower point. These steps are called *reflections* and are constructed in such a way that they maintain a finite inner n-dimensional volume of the simplex. Hence, the method expands the simplex in one or another direction to take larger steps. When it reaches a "valley floor", the

method leads to contraction in the transverse direction and tries to ooze down the valley. If there is a situation where the simplex must shrink in all directions, the method leads to contraction in all directions, pulling itself around its lowest point. In figure 5.1, the basic moves of this method are illustrated.

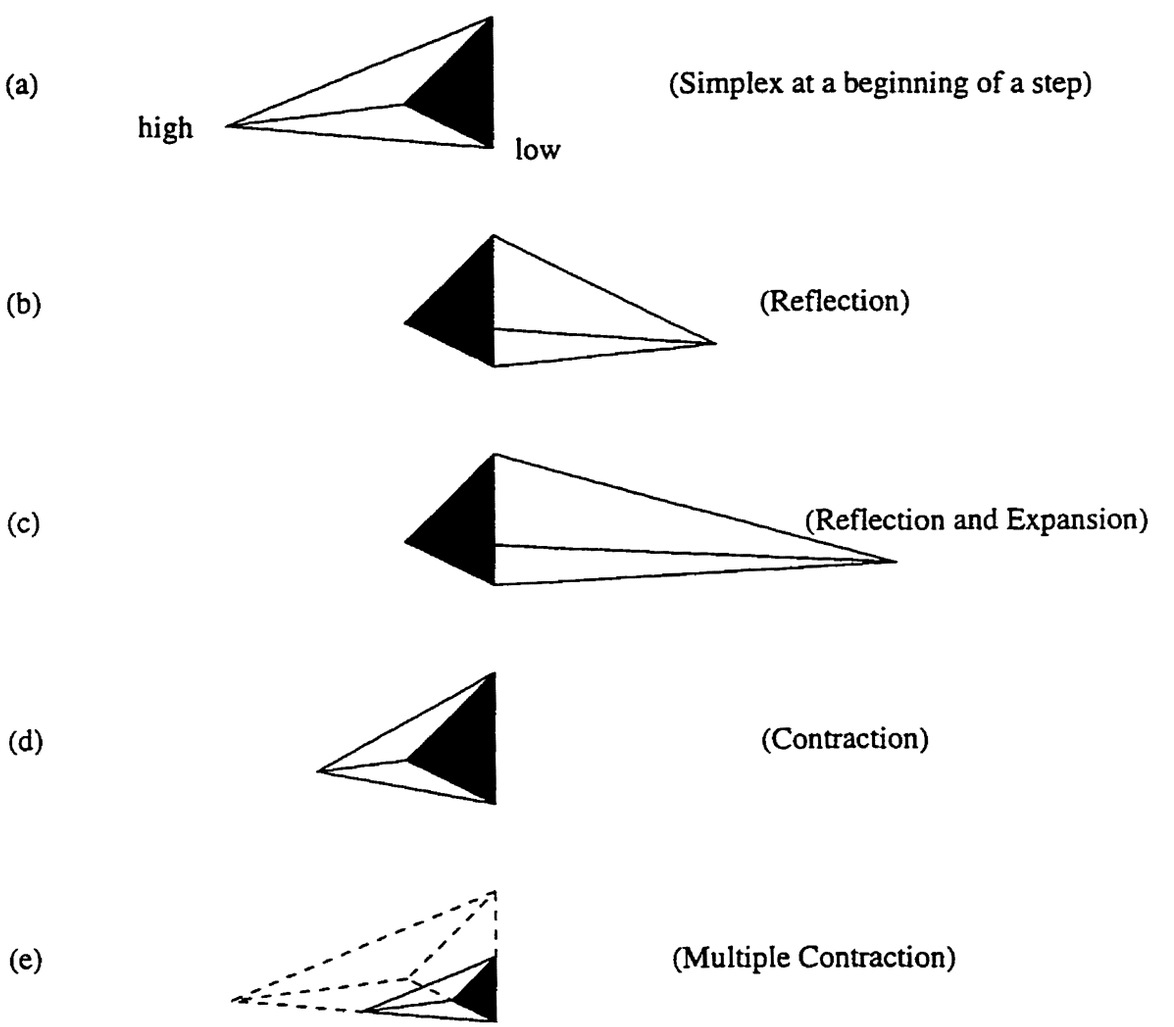


Figure 5.1: The basic moves of the simplex method

To present mathematical expressions for reflection, contraction, and expansion, we consider $P_0, P_1, \cdots, P_n$ as the $(n+1)$ points defining the current simplex. We

write y_i for the function value at point P_i and define

$$y_h = \max_i(y_i) \tag{5.4}$$

$$y_l = \min_i(y_i) \tag{5.5}$$

Accordingly, reflection, contraction and expansion are respectively defined by the following relations:

$$P^* = (1+\overline{\alpha})\overline{P} - \overline{\alpha}P_h \tag{5.6}$$

$$P^* = (1-\overline{\beta})\overline{P} - \overline{\beta}P_h \tag{5.7}$$

$$P^* = (1-\overline{\gamma})\overline{P} - \overline{\gamma}P_h \tag{5.8}$$

here P_h, P_l are the points which correspond to y_h and y_l; $\overline{P}$ is the centroid of the points with $i \neq h$; $\overline{\alpha}$, the reflection coefficient, is a positive constant; $\overline{\beta}$, the contraction coefficient, lies between 0 and 1; and $\overline{\gamma}$, the expansion coefficient, is greater than unity. An appropriate sequence of such moves will eventually bring all points into a valley.

Using this simplex method, we can summarize the ray tracing technique as shown in Figure 5.2. This technique has been encoded and used successfully in ray tracing in two dimensional stress induced heterogeneous anisotropic media.

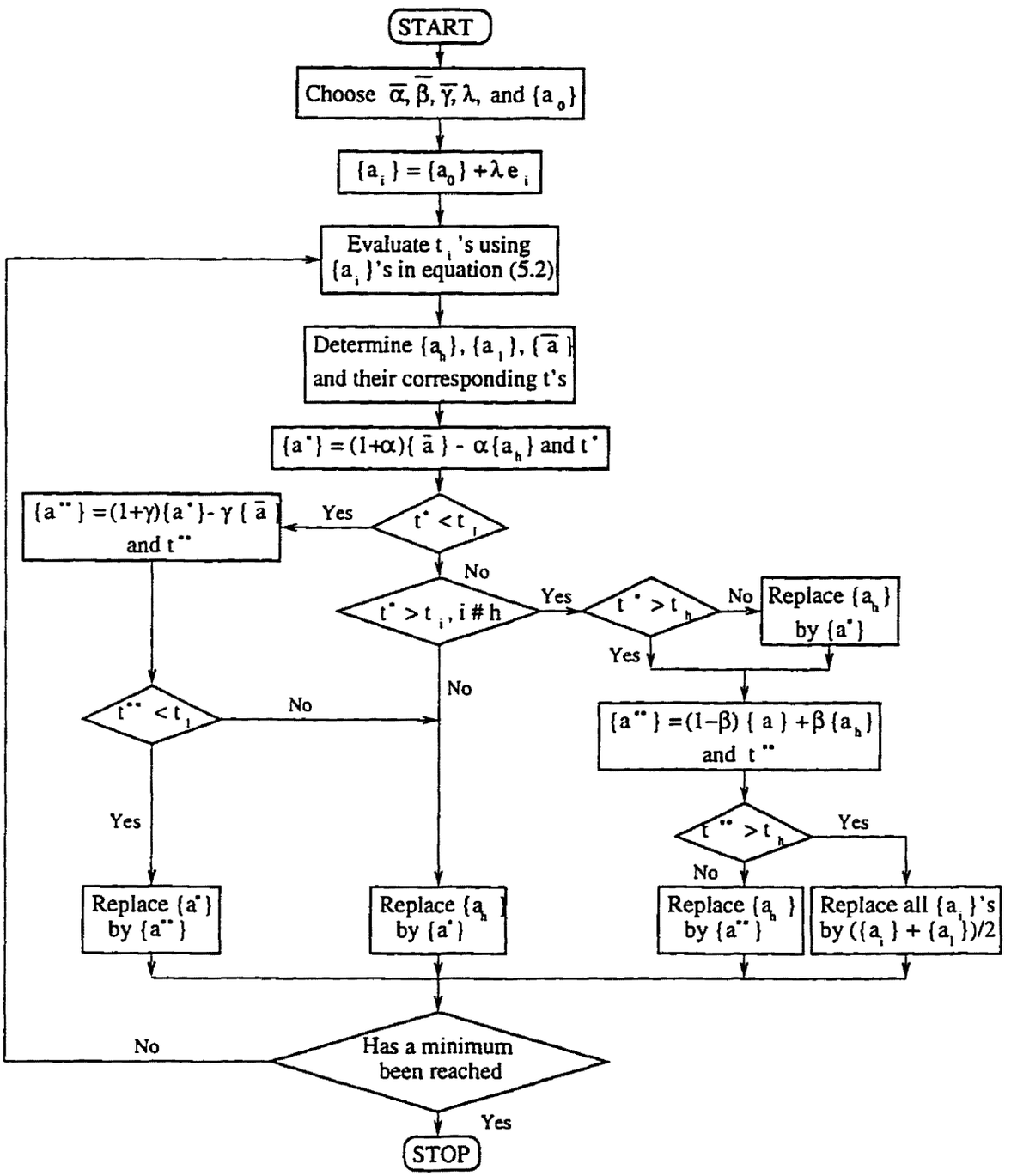


Figure 5.2: Flow diagram of the ray tracing procedure

5.3 Numerical Formulation of the Inverse Problem

In Chapter 4, based on the nodal stress approach and the stress potential function approach, we presented two mathematical models for formulating the inverse problem. The first approach led to a large scale nonlinear least-squares problems, while the second approach led to a constrained nonlinear least-squares problem. The main governing equation in both of these approaches has the form of a Fredholm integral equation of the first kind, except that they are nonlinear.

From the inverse problem perspective, Fredholm integral equations of the first kind lack the three requirements: existence, uniqueness, and stability of well-posedness [Gro93]. Therefore, our inversion algorithm should be capable of encountering an ill-posed and yet nonlinear system. In this section, after we set up the governing equations in a manageable form, we describe iterative schemes for solving these problems.

5.3.1 Nodal Stress Approach: Numerical Formulation

In this approach, the soil medium is discretized to a set of nodes, where each node carries three unknowns (in 2-D cases). As shown in Chapter 4, this approach leads to the travel-time equation (4.23); this leads to a set of nonlinear equations which is definitely inadequate for solving the inverse problem. There, we introduced the equilibrium equations and a regularization technique to remove the singularity of the system and to avoid undesired solutions. The combination of the travel-time equations, the equilibrium equations, and the regularization leads to the minimization problem (4.30) which we solve using an iterative optimization method.

The Governing Equations

The first governing equation is equation(4.23) in which σ_x, σ_y, and τ represent stress components at points lying on the ray path defined by $R = y$. These points are not necessarily nodal points. Therefore, we use an interpolation function to relate stress components inside a rectangular cell in the discretized medium to nodal components of the cell. Figure 5.3 shows a rectangular cell with its nodal points.

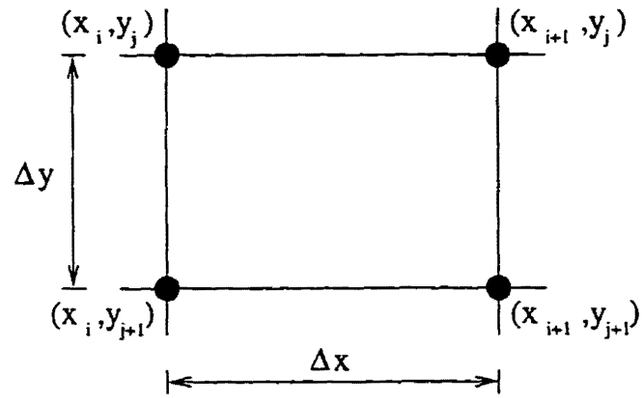


Figure 5.3: A rectangular cell with nodal points in a 2-D medium

The interpolation function for σ_x is defined in a compact form as follows:

$$\sigma_x(x,y) = \sum_{k=1}^{4} N_k \sigma_{xk} \tag{5.9}$$

where

$$\begin{aligned}
N_1 &= \tfrac{1}{\Delta x \Delta y}(x_{i+1} - x)(y_{j+1} - y) &&, \sigma_{x1} = \sigma_x(i,j) \\
N_2 &= \tfrac{1}{\Delta x \Delta y}(x - x_i)(y_{j+1} - y) &&, \sigma_{x2} = \sigma_x(i+1,j) \\
N_3 &= \tfrac{1}{\Delta x \Delta y}(x_{i+1} - x)(y - y_j) &&, \sigma_{x3} = \sigma_x(i,j+1)
\end{aligned}$$

and

$$N_4 = \tfrac{1}{\Delta x \Delta y}(x - x_i)(y - y_j) \qquad .\sigma_{x4} = \sigma_x(i+1,j+1)$$

Thus, to convert the travel-time integral equation (4.23) to a manageable algebraic equation, we introduce f_i as follows:

$$f_i = \sqrt{[1 + (\frac{dy}{dx})_i^2 - g_i]\sigma_{1i}^{-2\beta} + [1 + (\frac{dy}{dx})_i^2 + g_i]\sigma_{2i}^{-2\beta}} \tag{5.10}$$

where g_i, σ_{1i} and σ_{2i} can be obtained, respectively, from equations (4.24) , (4.19) and (4.20) in which σ_x, σ_y and τ must be obtained from equation (5.9). Thus using Simpsons rule, we obtain the following algebraic travel-time equation:

$$t = \frac{\Delta x}{3\sqrt{2}\alpha}[f_1 + 4f_2 + 2f_3 + 4f_4 + \cdots + 2f_{N-2} + 4f_{N-1} + f_N] \tag{5.11}$$

or in a compact form:

$$t = \frac{\Delta x}{3\sqrt{2}\alpha}\sum_{i=1}^{N}\omega_i f_i \tag{5.12}$$

where ω_i denotes the factors $1, 4, 2, 4, \cdots, 2, 4, 1$ and N is the number of discrete points we choose arbitrarily along the ray path.

To obtain algebraic equations corresponding to the equilibrium equations (4.25) and (4.26), we can either substitute the derivatives in these equations by appropriate finite difference expressions, or write the force equilibrium equations in x and y directions for a cell assuming linear variations for stress components; this assumption is consistent with the interpolation function (5.9).

Figure 5.4 shows a cell under stress components with linear variations. Considering a rectangular cell that undergoes these stress components simultaneously leads to the following equation corresponding to $\sum F_x = 0$ and $\sum F_y = 0$, respectively:

$$\begin{aligned}&\{(\sigma_x(i,j) + \sigma_x(i,j+1)) - (\sigma_x(i+1,j) + \sigma_x(i+1,j+1))\}\Delta y \\ &-\{(\tau(i,j) + \tau(i+1,j)) - (\tau(i,j+1) + \tau(i+1,j+1))\}\Delta x \;=\; 0\end{aligned} \tag{5.13}$$

$$\begin{aligned}&\{(\sigma_y(i,j) + \sigma_y(i+1,j)) - (\sigma_y(i,j+1) + \sigma_y(i+1,j+1))\}\Delta x \\ &-\{(\tau(i,j) + \tau(i,j+1)) - (\tau(i,j+1) + \tau(i+1,j+1))\}\Delta y + \gamma_s \Delta x \Delta y \;=\; 0\end{aligned} \tag{5.14}$$

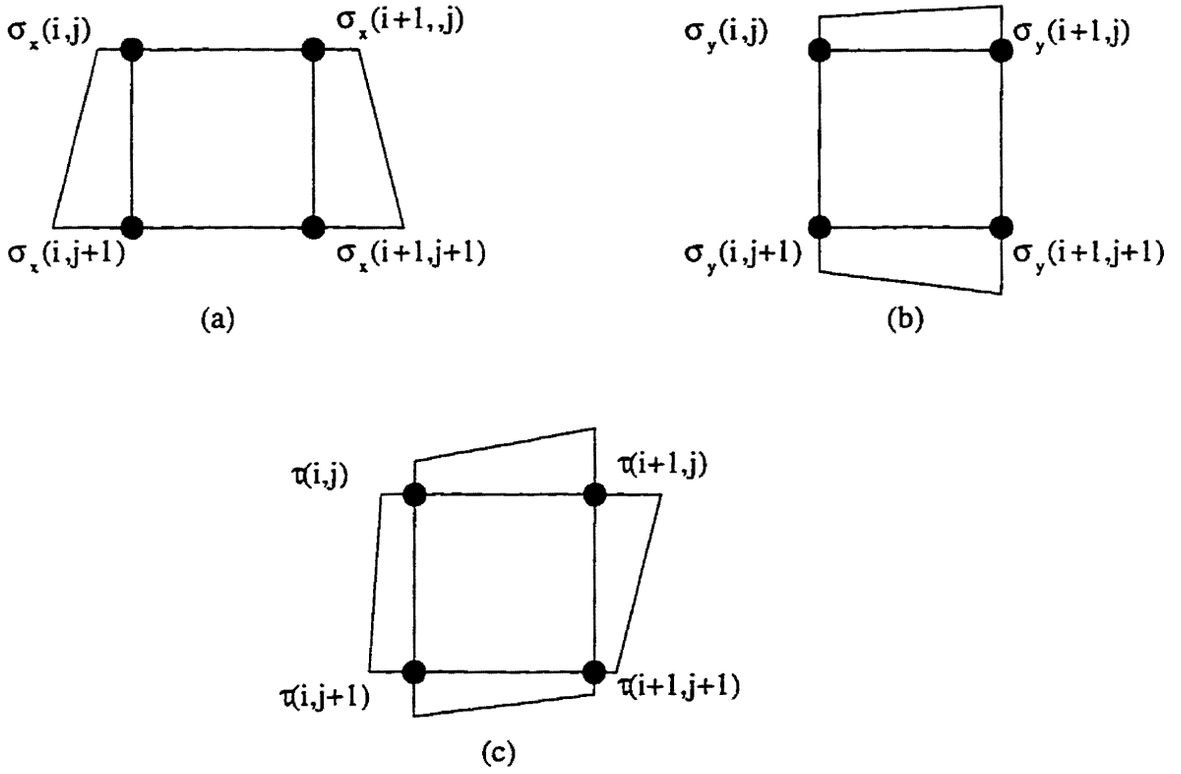


Figure 5.4: A rectangular cell with linear variations under: (a) σ_x, (b) σ_y, (c) τ

For regularization, we slightly modify the Laplacian operator used in [LD80, Iva87] for constructing the C matrix in (4.27). In 2-D parameter distribution, the Laplacian of a parameter σ is defined as follows:

$$\Delta\sigma = \frac{\partial^2\sigma}{\partial x^2} + \frac{\partial^2\sigma}{\partial y^2} \tag{5.15}$$

When used in regularization, this operator forces identical degrees of smoothness in the x and y directions; this is not necessarily desirable. To obtain a more flexible operator, we introduce the following operator:

$$\Delta\sigma = \lambda\frac{\partial^2\sigma}{\partial x^2} + \mu\frac{\partial^2\sigma}{\partial y^2} \tag{5.16}$$

where λ and μ, the smoothing parameters, must be estimated properly. If we choose $\lambda = \mu$, we can use the widely used L-curve method [Han92, HO93] for estimating an appropriate λ. If we choose to have two different parameters, we can use the idea of a solution roughness trade-off curves used in [CP92, PC92].

With the finite difference approximations, equation (5.16) leads to the following equation:

$$\begin{aligned}(\Delta\sigma)_{ij} \;=\; & \frac{\lambda}{(\Delta x)^2}(\sigma(i+1,j)+\sigma(i-1,j)) \\ & -2\left(\frac{\lambda}{(\Delta x)^2}+\frac{\mu}{(\Delta y)^2}\right)(\sigma(i,j)+\frac{\mu}{(\Delta y)^2}(\sigma(i+1,j)+\sigma(i,j-1)\end{aligned} \tag{5.17}$$

where the σ's represent the nodal parameter values in the neighborhood of node (i,j) as shown in Figure 5.5. Equation (5.17) must be applied to all interior nodes in the region; for nodes on the boundaries of the region, one-dimensional second derivative-based operators can be used. The different relations obtained in this way can be assembled properly to form the C matrix.

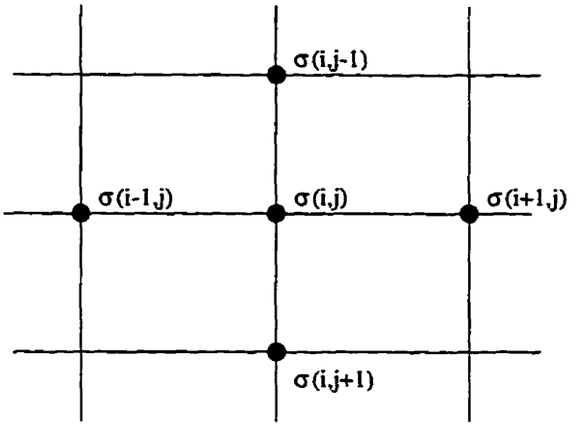


Figure 5.5: Nodal parameters used in forming the regularization operator

Nonlinear Least Squares Problem

For each measured travel-time, equation (5.12) provides a nonlinear equation, thus we obtain a set of nonlinear travel-time equations. Equations (5.13) and (5.14) can be applied to each cell in the discretized soil medium to produce a set of equations. We represent the set of nonlinear equations and the set of linear equations, respectively, by:

$$R_N(\sigma) = 0 \tag{5.18}$$

$$R_L(\sigma) = 0 \tag{5.19}$$

Therefore, after forming the matrix C, the regularization matrix, we can formulate the solution to the inverse problem as the solution of the following nonlinear least-squares problem:

$$\text{Minimize } \frac{1}{2}\left\| \begin{array}{c} R_N(\sigma) \\ R_L(\sigma) \\ C\sigma \end{array} \right\|^2 = \frac{1}{2}\|R(\sigma)\|^2 \tag{5.20}$$

Because of the large number of unknowns usually involved in this least-squares problem, and on the other since each governing equation involves only a small number of unknowns, this problem may be regarded as a large-scale sparse nonlinear least-squares problem. To solve this problem, we use the algorithm introduced by Martinez [Mar87]. In Chapter 3, we presented a step by step version of this algorithm in which Jacobian matrix evaluation, and sparse matrix multiplication, sparse linear system solution, and a typical bidimensional minimization problem are involved. In the following, we will present some details related to these operations.

Jacobian Matrix

In general, the Jacobian matrix J is defined as follows:

$$\begin{cases} J &= [J_{ij}] \\ J_{ij} &= \frac{\partial F_i}{\partial x_j} \end{cases} \quad ,i = 1, \cdots, m \;\; ; j = 1, \cdots, n_p \tag{5.21}$$

Here, F_i is the ith component of the function vector, x_j is the jth variable, and m and n_p are the number of function components and nodes, respectively. Thus, assuming n_t and n_c as, respectively, the number of measured travel- times and number of cells, we can obtain the Jacobian matrix using the following equations: For $1 \leq i \leq n_t$, according to equation (5.12):

$$J_{ij} = \frac{\partial t_i}{\partial \sigma_{x_j}} = \frac{\Delta x}{3\sqrt{2}\alpha} \sum_{k=1}^{N} \omega_k \frac{\partial f_k}{\partial \sigma_{x_j}} \tag{5.22}$$

$$J_{ij} = \frac{\partial t_i}{\partial \sigma_{y_j}} = \frac{\Delta x}{3\sqrt{2}\alpha} \sum_{k=1}^{N} \omega_k \frac{\partial f_k}{\partial \sigma_{y_j}} \tag{5.23}$$

$$J_{ij} = \frac{\partial t_i}{\partial \tau_j} = \frac{\Delta x}{3\sqrt{2}\alpha} \sum_{k=1}^{N} \omega_k \frac{\partial f_k}{\partial \tau_j} \tag{5.24}$$

To obtain appropriate expressions for $\frac{\partial f_k}{\partial \sigma_{x_j}}$, $\frac{\partial f_k}{\partial \sigma_{y_j}}$, $\frac{\partial f_k}{\partial \tau_j}$, we apply the chain rule,

thus

$$\begin{Bmatrix} \frac{\partial f_k}{\partial \sigma_{x_j}} \\ \frac{\partial f_k}{\partial \sigma_{y_j}} \\ \frac{\partial f_k}{\partial \tau_j} \end{Bmatrix} = \begin{bmatrix} \frac{\partial g_k}{\partial \sigma_{x_j}} & \frac{\partial \sigma_{1k}}{\partial \sigma_{x_j}} & \frac{\partial \sigma_{2k}}{\partial \sigma_{x_j}} \\ \frac{\partial g_k}{\partial \sigma_{y_j}} & \frac{\partial \sigma_{1k}}{\partial \sigma_{y_j}} & \frac{\partial \sigma_{2k}}{\partial \sigma_{y_j}} \\ \frac{\partial g_k}{\partial \tau_j} & \frac{\partial \sigma_{1k}}{\partial \tau_j} & \frac{\partial \sigma_{2k}}{\partial \tau_j} \end{bmatrix} \begin{Bmatrix} \frac{\partial f_k}{\partial g_k} \\ \frac{\partial f_k}{\partial \sigma_{1k}} \\ \frac{\partial f_k}{\partial \sigma_{2k}} \end{Bmatrix} \tag{5.25}$$

The partial derivatives of f_k w.r.t g_k, σ_{1k}, and σ_{2k} can be easily obtained from equation (5.10), while those of g_k, σ_{1k} and σ_{2k} w.r.t. σ_{x_j}, σ_{y_j} and τ_j can be evaluated from equations (4.24), (4.19) and (4.20), respectively.

To determine α and β, as described in Chapter 4, we treat them as unknowns; this leads to the following expressions for the corresponding Jacobian matrix members:

$$J_{ij} = \frac{\partial t_i}{\partial \alpha} = \frac{-\Delta x}{3\sqrt{2}\alpha^2} \sum_{k=1}^{N} \omega_k f_k \tag{5.26}$$

and

$$J_{ij} = \frac{\partial t_i}{\partial \beta} = \frac{-\beta \Delta x}{3\sqrt{2}\alpha} \sum_{k=1}^{N} \frac{\omega_k}{f_k} ([\{1 + (\frac{dy}{dx})_k - g_k]\sigma_{1k}^{-(2\beta+1)} + [\{1 + (\frac{dy}{dx})_k + g_k]\sigma_{2k}^{-(2\beta+1)}) \tag{5.27}$$

For $n_t + 1 \leq i \leq n_t + n_c$, according to equation (5.13):

$$J_{ij} = \frac{\partial R_{L_i}}{\partial \sigma_{x_j}} = \mp \Delta y \tag{5.28}$$

where the positive sign is used when j refers to points on the left side of the cell, and the negative sign for the right side.

$$J_{ij} = \frac{\partial R_{L_i}}{\partial \sigma_{y_j}} = 0 \tag{5.29}$$

and

$$J_{ij} = \frac{\partial R_{L_i}}{\partial \tau_j} = \mp \Delta x \tag{5.30}$$

where the positive and negative signs are respectively used when j refers to points on the bottom and top sides of the cell.

For $n_t + n_c + 1 \leq i \leq n_t + 2n_c$, according to equation (5.14):

$$J_{ij} = \frac{\partial R_{L_i}}{\partial \sigma_{x_j}} = 0 \tag{5.31}$$

$$J_{ij} = \frac{\partial R_{L_i}}{\partial \sigma_{y_j}} = \mp \Delta x \tag{5.32}$$

where the positive and negative signs are respectively used when j refers to points on the top side and bottom side of the cell.

$$J_{ij} = \frac{\partial R_{L_i}}{\partial \tau_j} = \mp \Delta y \tag{5.33}$$

where the positive and negative signs are respectively used when j corresponds to points on the right and the left sides of the cell.

For $n_t + 2n_c + 1 \leq i \leq n_t + 2n_c + 3n_p$, we have:

$$[J_{ij}] = [C_{ij}] \tag{5.34}$$

where C can be assembled according to equation (5.17).

Sparse Matrix Transpose and Multiplication

In step 2 of the iterative procedure of Martinez algorithm [Mar87], the following linear system must be solved in which a sparse matrix J_k represents the Jacobian matrix:

$$\| J_k^T J_k \omega_k + g_k \| \leq \eta_k \| g_k \| \tag{5.35}$$

To solve this system, we form the linear system

$$J_k^T J_k \omega_k = -g_k$$

and solve it iteratively using a sparse solver based on a preconditioned conjugate gradient method [BBC+94]; we stop iterations when (5.35) is satisfied. Since J_k^T

is sparse and usually a large matrix, the calculation of J_k^T and $J_k^T J_k$ required for forming the linear system must be carried out using an efficient method. For this purpose we use the efficient algorithms invented by Gustavson [Gus78] for obtaining a sparse matrix transpose and sparse matrix multiplications.

Bidimensional Minimization Problem

In steps 3 and 5 of the Martinez algorithm, a typical bidimensional minimization problem of the following form must be solved:

$$\text{Minimize } \|J_k d + R_k\|$$
$$\text{subject to } \begin{cases} d = \lambda_1 d_k^1 + \lambda_2 d_k^2 \\ \|d\| \leq t \end{cases}$$

Here J_k is a given matrix, R_k, d_k^1 and d_k^2 are given vectors, and t is a given scalar. Thus our aim is to find the coefficients λ_1 and λ_2. To formulate a solution to this problem, we first introduce $E = \langle e_1, e_2 \rangle$ to be an orthonormal basis of the subspace spanned by $\langle d_k^1, d_k^2 \rangle$; we calculate E using singular values decomposition [PTVF92]. Thus, the above problem can be written as:

$$\text{Minimize } \|J_k d + R_k\|$$
$$\text{subject to } \begin{cases} d = E\mu \\ \|d\| \leq t \end{cases} \tag{5.36}$$

where $\mu^T = \langle \mu_1, \mu_2 \rangle$ and μ_1 and μ_2 are unknown coefficients. Since $d = E\mu$, this problem can be equivalently written as:

$$\text{Minimize } \|J_k E\mu + R_k\|^2$$
$$\text{subject to } \|\mu\|^2 \leq t^2$$

or

$$\text{Minimize } (\mu^T B\mu + 2\mu^T C + D)$$
$$\text{subject to } \mu^T \mu \leq t^2$$

where B, C, and D are defined as follows:

$$B = \begin{bmatrix} b_{11} & b_{22} \\ b_{21} & b_{22} \end{bmatrix} = E^T J_k^T J_k E \tag{5.37}$$

$$C = \begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = E^T J_k^T R_k \tag{5.38}$$

and

$$D = R_k^T R_k \tag{5.39}$$

To solve this bidimensional minimization problem, we first assume that the constraint is inactive. Thus, for the quadratic function in (5.36) to be stationary, we must have:

$$B\mu + C = 0 \tag{5.40}$$

If μ, the solution of this 2 by 2 linear system, satisfies the constraint, the correct solution is achieved; otherwise, the constraint is active and we form the following Lagrange function in which ρ is an unknown Lagrange multiplier:

$$L(\mu, \rho) = \mu^T B\mu + 2\mu^T C + D + \rho(\mu^T \mu - t^2) \tag{5.41}$$

The necessary conditions for a stationary point for this function are:

$$\frac{\partial L}{\partial \mu} = (B + \rho I)\mu + C = 0 \tag{5.42}$$

$$\frac{\partial L}{\partial \rho} = \mu^T \mu - t^2 = 0 \tag{5.43}$$

This is a 3 by 3 system of nonlinear equations that we solve iteratively using Newton-Raphson's method.

Flow Chart of the Inverse Problem

The numerical procedure of solving the inverse problem using the nodal stress approach is summarized in Figure 5.6. Based on this procedure, we have developed a FORTRAN code to be used in stress reconstruction of 2-D soil media.

5.3.2 Stress Function Approach: Numerical Formulation

As described in Chapter 4, in this approach, only one continuous function is required to represent the distribution of the three stress components in 2-D problems. Here, we chose polynomials as stress potential functions; Hence, the inverse problem solution is a matter of determining the unknown coefficients of a polynomial.

The only governing equation in this approach is the travel-time equation (4.23) where σ_1, σ_2 and g must be obtained by substituting σ_x, σ_y and τ from equations (4.35) to (4.37) in equations (4.19), (4.20), and (4.24) respectively. As with the previous approach, we obtain equation (5.12), the algebraic version of equation (4.23). With the constraints represented by (4.38), (4.40), (4.41) and (4.42), the inverse problem leads to the constrained nonlinear least-squares problem represented by (4.45). We solve this problem using an algorithm based on the SQP algorithm of Scittkowski [Sch88]. As described in Chapter 4, this algorithm mainly involves calculating the Jacobian matrix of the constraints and the functions values. We next present relations for evaluating the Jacobian matrices.

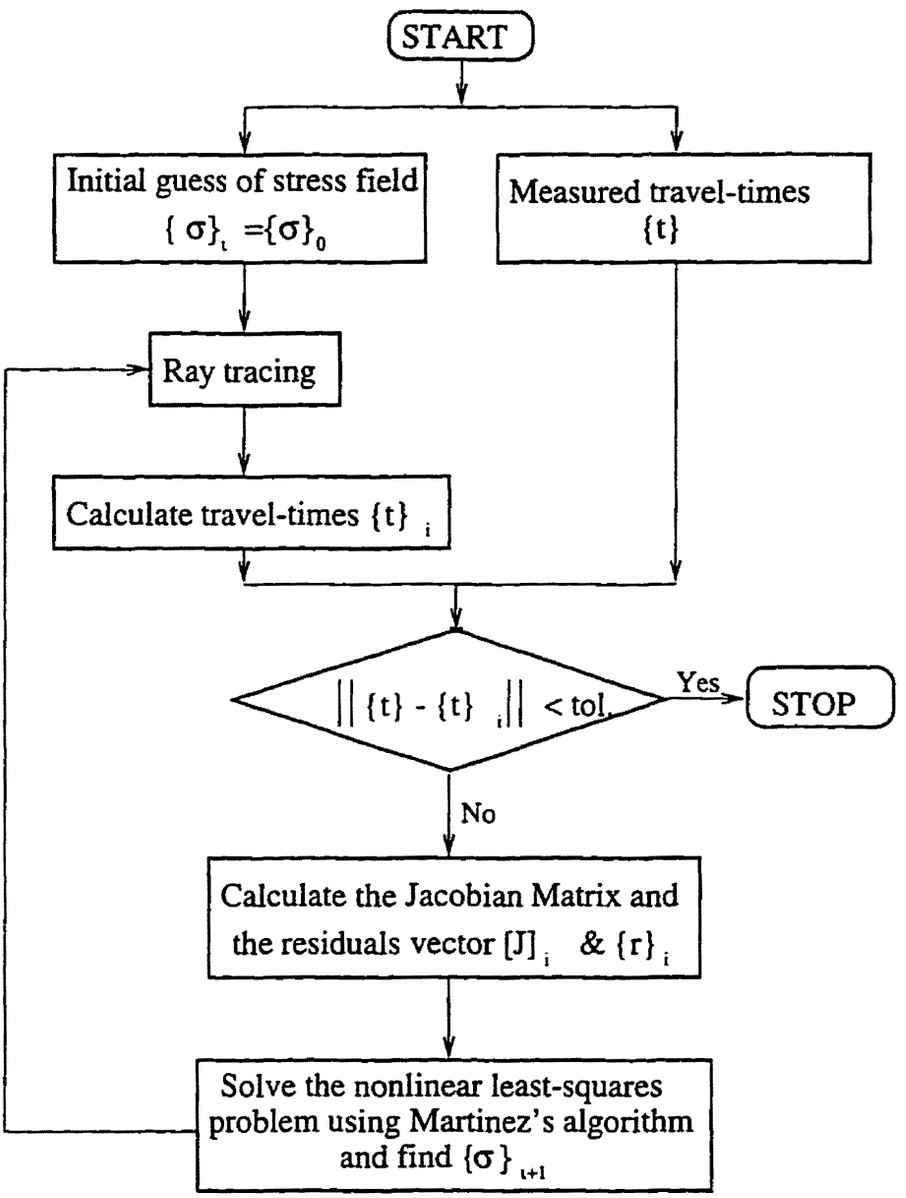


Figure 5.6: Flow Chart of the numerical solution in the nodal stress approach

Jacobian Matrix of the Travel-Time Equations

With $(a_j, j = 1, \cdots, n)$ as the unknown coefficients, and n_t as the number of measured travel-times, according to equation (ref{eq:5.12):

$$J_{ij} = \frac{\partial t_i}{\partial a_j} = \frac{\Delta x}{3\sqrt{2}\alpha} \sum_{k=1}^{N} \omega_k \frac{\partial f_k}{\partial a_j} \quad , \text{For } 1 \leq i \leq n_t \tag{5.44}$$

where $\frac{\partial f_k}{\partial a_j}$ can be calculated using the following chain rule:

$$\frac{\partial f_k}{\partial a_j} = \langle \frac{\partial f_k}{\partial g_k} \; \frac{\partial f_k}{\partial \sigma_{1k}} \; \frac{\partial f_k}{\partial \sigma_{2k}} \rangle \begin{bmatrix} \frac{\partial g_k}{\partial \sigma_{x_k}} & \frac{\partial g_k}{\partial \sigma_{y_k}} & \frac{\partial g_k}{\partial \tau_k} \\ \frac{\partial \sigma_{1k}}{\partial \sigma_{x_k}} & \frac{\partial \sigma_{1k}}{\partial \sigma_{y_k}} & \frac{\partial \sigma_{1k}}{\partial \tau_k} \\ \frac{\partial \sigma_{2k}}{\partial \sigma_{x_k}} & \frac{\partial \sigma_{2k}}{\partial \sigma_{y_k}} & \frac{\partial \sigma_{2k}}{\partial \tau_k} \end{bmatrix} \begin{Bmatrix} \frac{\partial \sigma_{x_k}}{\partial a_j} \\ \frac{\partial \sigma_{y_k}}{\partial a_j} \\ \frac{\partial \tau_k}{\partial a_j} \end{Bmatrix} \tag{5.45}$$

The partial derivatives of f_k w.r.t g_k, σ_{1k}, and σ_{2k} can be easily obtained from equation (5.10); those of g_k, σ_{1k} and σ_{2k} w.r.t. σ_{x_k}, σ_{y_k} and τ_k must be evaluated from equations (4.24), (4.19) and (4.20), respectively; and finally those of σ_{x_k} σ_{y_k} and τ w.r.t. a_j must be calculated using the stress functions such as equations (4.35) to (4.37).

To determine α and β, we need:

$$J_{ij} = \frac{\partial t_i}{\partial \alpha}$$

and

$$J_{ij} = \frac{\partial t_i}{\partial \beta}$$

These derivatives can be evaluated using equations (5.26) and (5.27).

Jacobian Matrix of the Constraints

In Chapter 4, we obtained two sets of constraints in this approach: linear constraints and nonlinear constraints. In computing the Jacobian matrix of the constraints,

we can treat both of these sets in a similar manner, except that the part of the Jacobian matrix corresponding to the linear constraints is constant, and thus must be computed once, while the rest must be updated in each iteration. The constant part of J_g, the constraints Jacobian matrix, can be obtained as follows:

$$(J_g)_{ij} = \frac{\partial \sigma_{x_i}}{\partial a_j}, \frac{\partial \sigma_{y_i}}{\partial a_j}, \text{ or } \frac{\partial \tau}{\partial a_j}, \; i = 1, \cdots, n_L \tag{5.46}$$

$$(J_g)_{ij} = \frac{\partial \sigma_{x_i}}{\partial a_j}, \; i = n_L + 1, \cdots, n_L + n_p \tag{5.47}$$

$$(J_g)_{ij} = \frac{\partial \sigma_{y_i}}{\partial a_j}, \; i = n_L + n_p + 1, \cdots, n_L + 2n_p \tag{5.48}$$

In equation (5.46), depending on the boundary conditions, some or all of the partial derivative terms may be used. The rest of J_g, which corresponds to the nonlinear constraints given by (4.42), can be calculated using the chain rule as follows:

$$(J_g)_{ij} = \sigma_{y_i} \frac{\partial \sigma_{x_i}}{\partial a_j} + \sigma_{x_i} \frac{\partial \sigma_{y_i}}{\partial a_j} - 2\tau \frac{\partial \tau}{\partial a_j}, \; i = n_L + 2n_p + 1, \cdots, n_L + 3n_p \tag{5.49}$$

In equations (5.46) to (5.49), the partial derivatives must be calculated using the functions of the corresponding stress component.

Flow Chart of The Inverse Problem

The numerical procedure of solving the inverse problem using the stress potential function approach is summarized in Figure 5.7. Based on this procedure, we have developed a FORTRAN code to be used in stress reconstruction of 2-D soil media.

5.4 Scaling

In the two approaches we presented for soil stress reconstruction, the unknown variables involved may vary greatly in magnitude. In the nodal stress approach, there

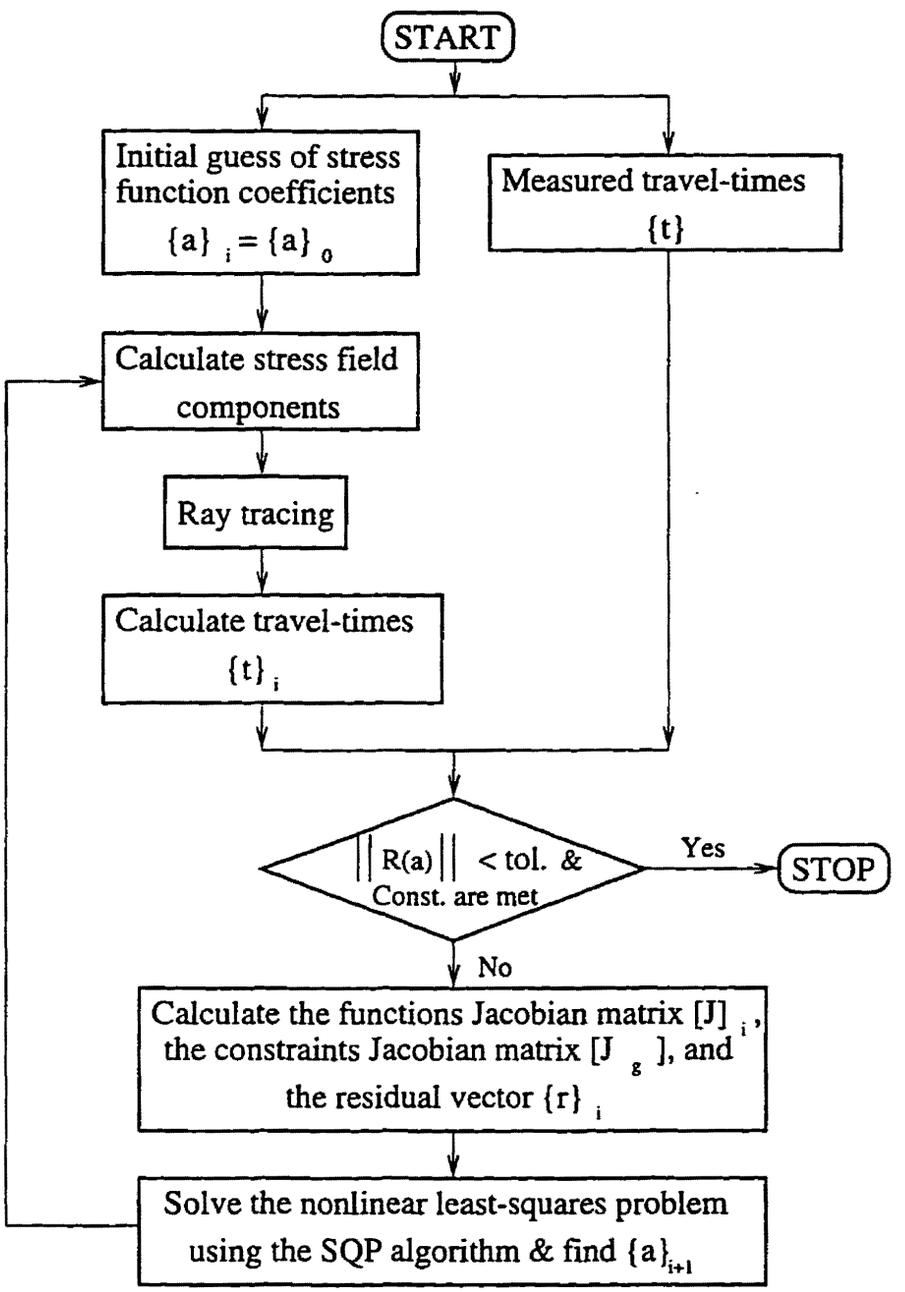


Figure 5.7: Flow Chart of the numerical solution in the stress potential function approach

are three unknown variables: two principal stresses and the angle these stresses make with the Cartesian system axes. The angle is in the range $[0, 2\pi]$, while the principal stresses, depending on their units, may be in a much larger range. In the stress potential function approach, with α and β as unknown variables, we have the same kind of difference between the magnitude of stress components, and α and β.

Large variations in the magnitude of variables lead to large variations in the magnitude of the corresponding derivatives. Therefore, optimization algorithms based on functions and constraints derivatives lead to ill-conditioned Jacobian and Hessian matrices; this, due to the finite precision available in computers, disturbs the convergence of the algorithms significantly. The remedy to this problem is re-scaling the independent variables; this is equivalent to changing their units.

For re-scaling, in case a realistic guess for lower and upper bound of each independent variable is available, Gill et al [GMW81] recommended the following transformation relation:

$$x = Dy + C \tag{5.50}$$

where $\{x_j\}$ are the original variables, $\{y_j\}$ are the transformed variables, and D and C are, respectively, a diagonal matrix and a vector defined in terms of the variables lower and upper bounds. For variable x_j with lower and upper bound a_j and b_j, the jth diagonal element of D is:

$$d_{jj} = \frac{1}{2}(b_j - a_j) \tag{5.51}$$

and the jth element of the vector C is:

$$C_j = \frac{1}{2}(b_j + a_j) \tag{5.52}$$

This transformation guarantees that $-1 \leq y_j \leq 1$ for j. If a realistic upper and lower bounds are not available, the most commonly used transformation of the

following form is recommended [GMW81, DS83]:

$$x = Dy \tag{5.53}$$

where D is a constant diagonal matrix. The element of this diagonal matrix must be chosen according to the typical values in such a way that the transformed variables vary, roughly, in the same range. In this work, we used a scaling scheme based on the latter transformation form.

An interesting aspect of scaling strategies based on (5.53) is that when they are used in Newton methods with exact Hessian matrix, they do not affect the sequence of iterates [GMW81, DS83]. This means that

$$x_k = Dy_k \tag{5.54}$$

where k is the iteration number.

Chapter 6

Numerical Examples

6.1 Introduction

In order to assess the performance of the mathematical models presented for ray tracing and stress reconstruction in particulate soil media, we have developed FORTRAN computer programs based on these models. In this chapter. we present the results we have obtained using these programs in some typical examples.

For ray tracing, we will present an example of a soil medium that undergoes a footing loading. First, we calculate the stress distribution corresponding to the loading, we then use the ray tracing program for determining the ray paths and their corresponding travel-times.

For stress reconstruction, we will present four typical examples in which synthetic data is used to infer the corresponding stress fields. In the first example where we presen a footing problem, we will reconstruct the stress fields using a program based on the nodal stress approach; in the second to fourth examples, where we, respectively, present footing, tunnel, and retaining wall problems, we

will use another program based on the stress potential function approach for stress reconstruction.

In the stress reconstruction examples, in order to obtain the travel-time data synthetically, we first use ABAQUS [HS95], a finite element software widely used in nonlinear problems, to analyze the soil medium. Having the stress distribution, we use the ray tracing program for generating the synthetic travel-time data.

6.2 Ray Tracing in a Soil Medium Under Footing

To illustrate the performance of the ray tracing technique we developed for stress induced anisotropic soil media, we present an example corresponding to a typical problem in soil mechanics. In this example, as depicted in Figure 6.1, we assume a region of soil medium that undergoes a footing loading. We assume that the footing is long enough so that, for stress analysis, a plane strain model is satisfactory. We, further, assume that the region is symmetric about a vertical plane, and that it is of infinite horizontal extent.

For stress analysis using ABAQUS, we assumed that soil with density $\gamma_s = 1.85\text{gr/cm}^3$ behaves as elastic-plastic material. The elastic response is assumed to be linear and isotropic with Young's modulus $E = 2.1 \times 10^3\text{kg/cm}^2$ and Poisson's ratio $\nu = 0.3$. Yield is assumed to be governed by the Mohr-Coulomb surface, with a friction angle $\Phi = 30°$, and cohesion $C = 0.25\text{kg/cm}^2$. Unfortunately, Mohr-Coulomb yield criterion is unavailable in ABAQUS, however it is possible to convert Mohr-Coulomb parameters to those of Drucker-Prager which is available in ABAQUS. According to the ABAQUS user manual [HS95], in plane strain problems, the following formula can be used for calculating β and σ_C^o, Drucker-Prager

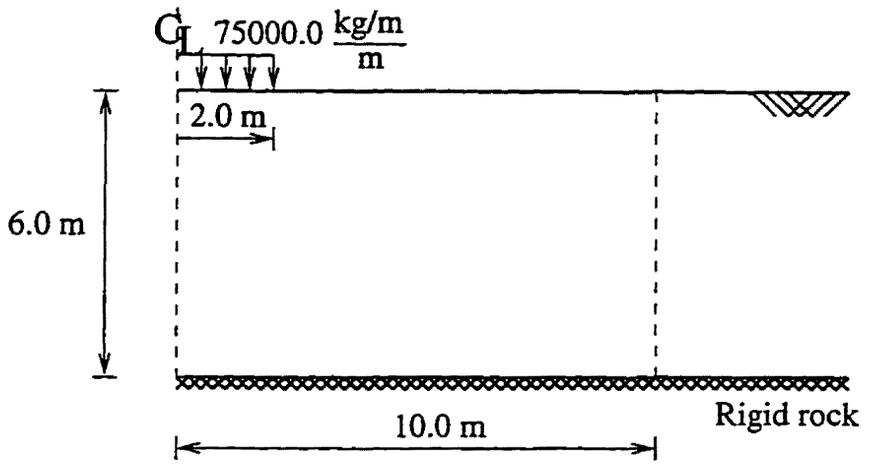


Figure 6.1: Soil region under footing

parameters corresponding to a given Φ and C:

$$\tan\beta = \frac{3\sqrt{3}\tan\Phi}{\sqrt{9+12\tan^2\Phi}} \tag{6.1}$$

$$d = \frac{3\sqrt{3}C}{\sqrt{9+12\tan^2\Phi}} \tag{6.2}$$

and

$$\sigma_C^\circ = \frac{d}{(1-\frac{1}{3}\tan\beta)} \tag{6.3}$$

Therefore, we find $\beta = 39.76°$ and $\sigma_C^\circ = 0.5\text{kg/cm}^2$. Having analyzed the soil medium using ABAQUS, we obtain the stress distribution shown in Figure 6.2. In this figure the distribution of σ_x, σ_y and τ is shown.

With the stress distribution results obtained from ABAQUS and assuming $\alpha = 40$ and $\beta = 0.22$ corresponding to the stress in (kg/m^2), we used the ray tracing program we have developed to determine the ray paths and travel-times. We found that using only three a_i's, the perturbation terms in equation (4.14), is quite satisfactory. Therefore, we obtained the ray paths between source and receiver points located in two parallel boreholes on the left and right boundary of the

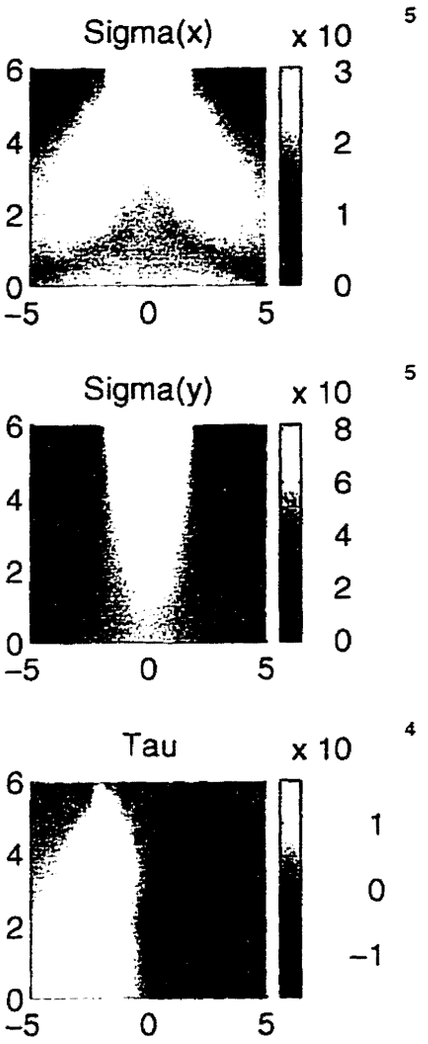


Figure 6.2: Distribution of σ_x, σ_y, and τ of the soil medium under footing

region respectively. These ray paths are shown in Figure 6.3. In this figure, the tendency of the ray paths to pass through the high stress zone is clearly shown.

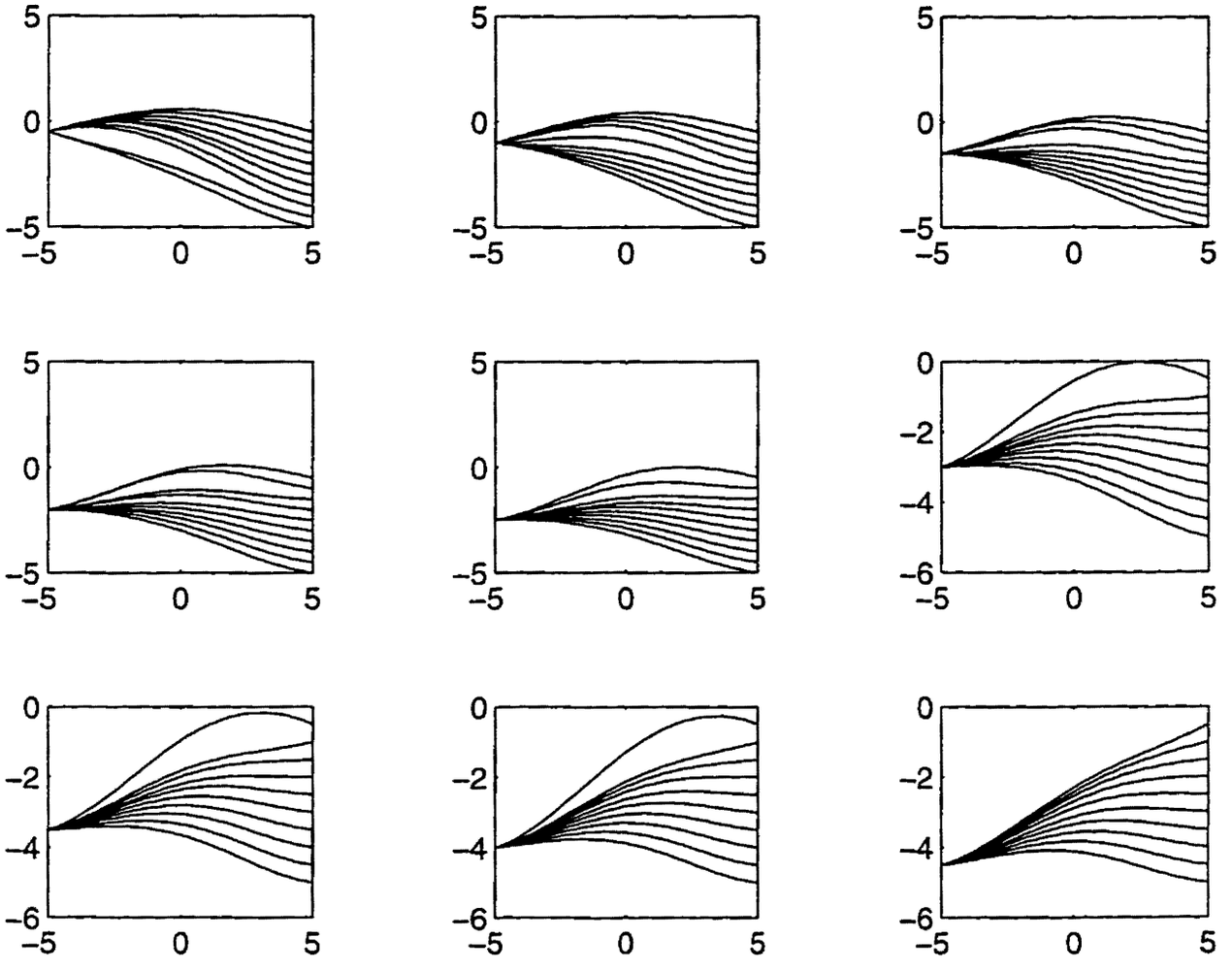


Figure 6.3: Ray paths between sources and receivers in a soil region under footing

6.3 Stress Reconstruction in Footing Problems

To demonstrate the performance of the stress reconstruction techniques we have developed in this work, we will present two footing problem examples. In the first example, we used the nodal stress approach to reconstruct the stress of a 25 node soil region; this approach failed to handle large-scale problems. In the second example, we used the stress potential function approach to reconstruct the stress field of the example presented in Section 6.2.

6.3.1 Footing Problem Stress Reconstruction: Nodal Stress Approach

In this example, we assume a soil region that undergoes a mat foundation of 10.0 m width. As depicted in Figure 6.4, the foundation causes a uniform pressure of 4000kg/m^2 on the contact region; the soil medium can be modeled as a half-space subjected to a uniform load.

Assuming that the soil in this example behaves as an elastic and isotropic medium, we can use the closed form solution of the stress distribution obtained using theory of elasticity [KS91]. However, in this solution, the gravity load of soil is not considered; to the author's knowledge, no analytical solution which accounts for the gravity load is available. To account for the gravity load, we superimpose the following approximate solution to the closed form solution:

$$\begin{cases} \sigma_x &= k\sigma_y \\ \sigma_y &= \gamma_s \sigma_y \\ \tau &= 0 \end{cases}$$

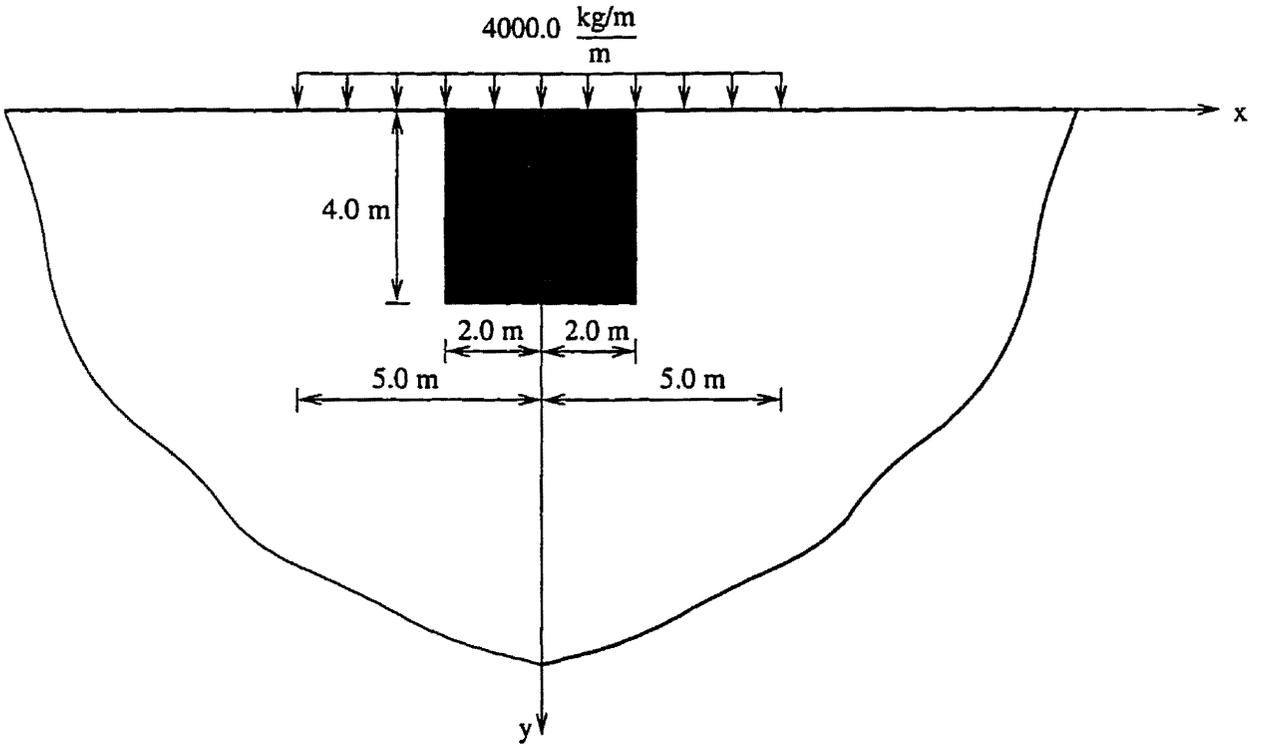


Figure 6.4: Soil region modeled as a half-space subjected to a uniform load

Here σ_x, σ_y and τ are the stress components, γ is the soil density, y is the depth of the point where the stress is evaluated, and k is a constant. Due to the limited capability of the nodal stress approach, we construct the stress distribution of a small subregion shown by the shaded area in Figure 6.4. Discretizing the shaded area into a 5×5 grid and choosing $k = 0.1$, we obtained the stress distribution in this subregion as shown in Figure 6.5.

With the calculated stresses, and $\alpha = 30$ and $\beta = 0.2$, we first generated travel-time data using the ray tracing algorithm. Then we input this travel-time data along with α and β to the FORTRAN program we have developed based on the nodal stress approach. The initial stress distribution supplied to the program is shown in Figure 6.6; the divergence of this stress field from the actual field is around %10. Finally we input the following equality constraints corresponding to points in contact with the footing:

$$\begin{cases} \sigma_y &= 4000\text{kg/m}^2 \\ \tau &= 0 \end{cases}$$

After eight iterations, the algorithm converges to the stress field shown in Figure 6.7.

Attempts to solve large-scale problems or even solving this problem but starting from an initial guess with more divergence have failed. Therefore, this approach is not appropriate for real world problems.

6.3.2 Footing Problem Stress Reconstruction: Stress Functions Approach

In this example, we attempt to reconstruct the stress field of the footing problem presented in Section 6.2. The travel-time data required for the stress reconstruction

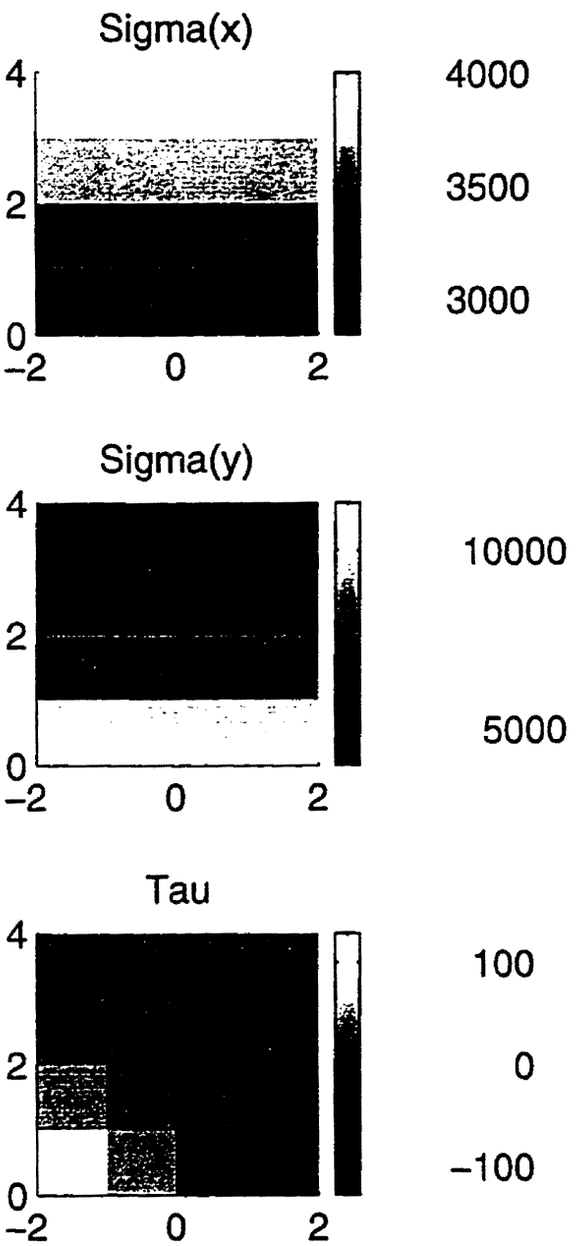


Figure 6.5: Distribution of σ_x, σ_y, and τ in the rectangular subregion

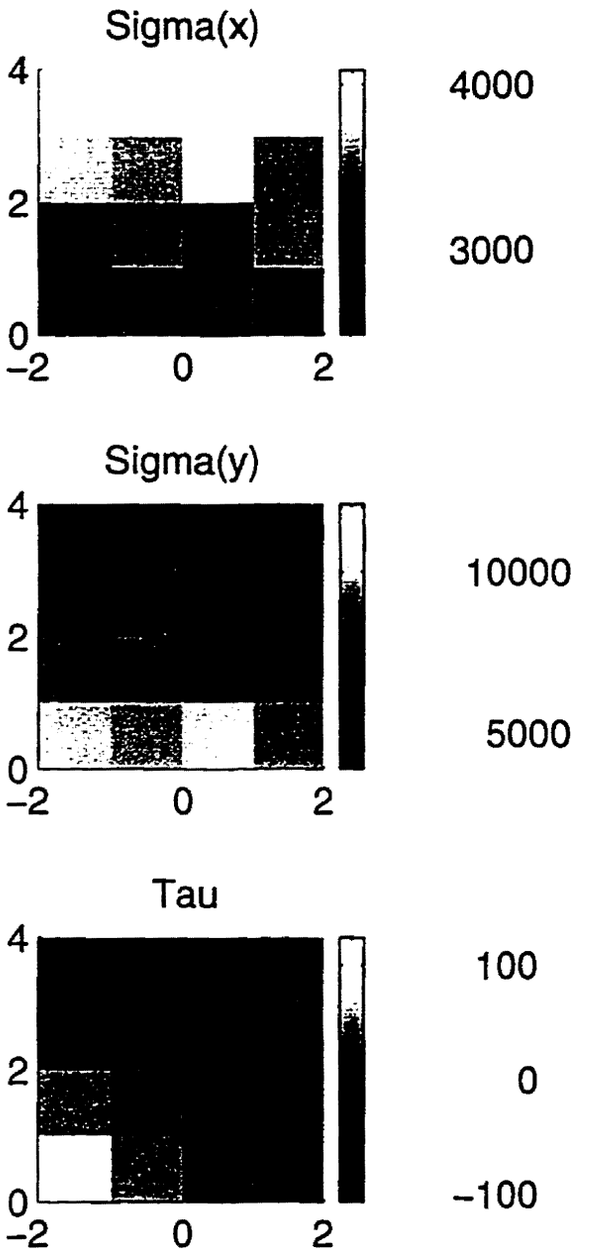


Figure 6.6: Distribution of σ_x, σ_y, and τ in the initial guess

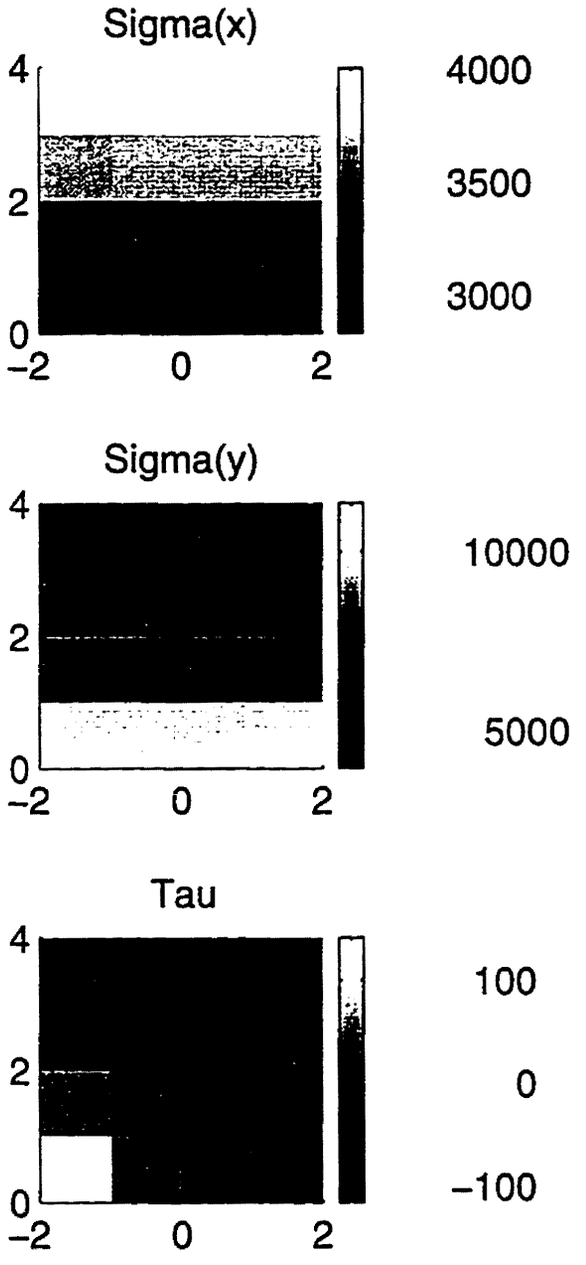


Figure 6.7: Reconstructed distribution of σ_x, σ_y, and τ in the footing problem using the nodal stress approach

was obtained from the ray tracing carried out in that section. As the loading and geometry are symmetric about the x-axis in the problem, we choose an even potential function of x. Furthermore, we found that polynomial of orders less than 6 do not present the stress field of this problem satisfactorily. Therefore, we use the following eighth order polynomial as the stress potential function:

$$\begin{aligned}\Phi &= a_1x^2 + a_2y^2 + a_3x^2y + a_4y^3 + a_5x^4 + a_6x^2y^2 \\ &\quad + a_7y^4 + a_8x^4y + a_9x^2y^3 + a_{10}y^5 + a_{11}x^6 + a_{12}x^4y^2 \\ &\quad + a_{13}x^2y^4 + a_{14}x^6y + a_{15}x^4y^3 + a_{16}x^2y^5 + a_{17}x^8 \\ &\quad + a_{18}x^6y^2 + a_{19}x^4y^4 \end{aligned} \tag{6.4}$$

This leads to the following relations for the stress field components:

$$\begin{aligned}\sigma_x &= \frac{\partial^2\Phi}{\partial y^2} + \gamma_s y \\ &= 2a_2 + (6a_4 + \gamma_s)y + 2a_6x^2 + 12a_7y^2 \\ &\quad + 6a_9x^2y + 20a_{10}y^3 + 2a_{12}x^4 + 12a_{13}x^2y^2 \\ &\quad + 6a_{15}x^4y + 20a_{16}x^2y^3 + 2a_{18}x^6 + 12a_{19}x^4y^2 \end{aligned} \tag{6.5}$$

$$\begin{aligned}\sigma_y &= \frac{\partial^2\Phi}{\partial x^2} + \gamma_s y \\ &= 2a_1 + (2a_3 + \gamma_s)y + 12a_5x^2 + 2a_6y^2 \\ &\quad + 12a_8x^2y + 2a_9y^3 + 30a_{11}x^4 + 12a_{12}x^2y^2 \\ &\quad + 2a_{13}y^4 + 30a_{14}x^4y + 12a_{15}x^2y^3 + 2a_{16}y^5 \\ &\quad + 56a_{17}x^6 + 30a_{18}x^4y^2 + 12a_{19}x^2y^4 \end{aligned} \tag{6.6}$$

$$\begin{aligned}\tau &= \frac{\partial^2\Phi}{\partial x\partial y} \\ &= 2a_3x + 4a_6xy + 4a_8x^3 + 6a_9xy^2 + 8a_{12}x^3y \\ &\quad + 8a_{13}xy^3 + 6a_{14}x^5 + 12a_{15}x^3y^2 + 10a_{16}xy^4 \\ &\quad + 12a_{18}x^5y + 16a_{19}x^3y^3 \end{aligned} \tag{6.7}$$

Thus in this approach we need to determine only 19 unknowns to reconstruct the stress field. In case α and β are unknown, we must determine 21 unknowns.

Boundary Conditions

On the surface of the soil region where $y = 0$, the stresses σ_y and τ are known. For points in contact with the footing ($-2.0 \leq x \leq 2.0$), we have:

$$\begin{aligned} \sigma_y &= 75000\,\mathrm{kg/m^2} \\ \tau &= 0 \end{aligned}$$

while for the free points on the surface ($2.0 \leq |x| \leq 5.0$), we have:

$$\begin{aligned} \sigma_y &= 0 \\ \tau &= 0 \end{aligned}$$

To incorporate these conditions and the linear and nonlinear constraints required in the reconstruction algorithm, we define a set of discrete points on the region. In this problem, we used the nodal points of a grid of 0.5m intervals in both x and y directions.

Reconstructed Stress Field

In this example and the upcoming examples, to demonstrate the results, we present four sets of three diagrams showing the distribution of σ_x, σ_y and τ; each set is placed in a column. The first, which we call the *actual*, is that computed by ABAQUS. The fourth, which we call the *best possible*, is the least squares approximation to the actual obtained by using the chosen polynomial form for the stress function. The second, which we call the *initial guess* is obtained by making random

perturbations to the coefficients in the best possible. The third, which we call the *reconstructed* field, is that obtained from the tomographic reconstruction.

In this example, the results are shown in Figure 6.8. The initial guess for stress distribution, shown in the second column of this figure, is obtained by making random perturbations, of amounts varying from %5 to %30, to the coefficients of the best possible. The initial guess of α and β is:

$$\alpha = 32.0 \quad \text{and} \quad \beta = 0.264.$$

Using the FORTRAN program we have developed based on the stress potential function approach, we obtained the reconstructed stress distributions shown in the third column of Figure 6.8 after three iterations, with

$$\alpha = 38.96 \quad \text{and} \quad \beta = 0.214.$$

These values are very close to those we used to generate the travel-time data. Between the constructed stress distributions shown in the third column and the best possible solution shown in the fourth column, there is a good agreement.

6.3.3 Tunnel Problem Stress Reconstruction: Stress Functions Approach

In this example, as depicted in Figure 6.9, we assume a symmetric region of a soil medium which encloses a tunnel. We assume that the tunnel is long enough so that a plane strain model is satisfactory.

To obtain the stress distribution of the soil medium using ABAQUS, we assumed that soil with density $\gamma_s = 1.85\text{gr/cm}^3$ behaves as elastic-plastic material. The elastic response is assumed to be linear and isotropic with Young's modulus $E =$

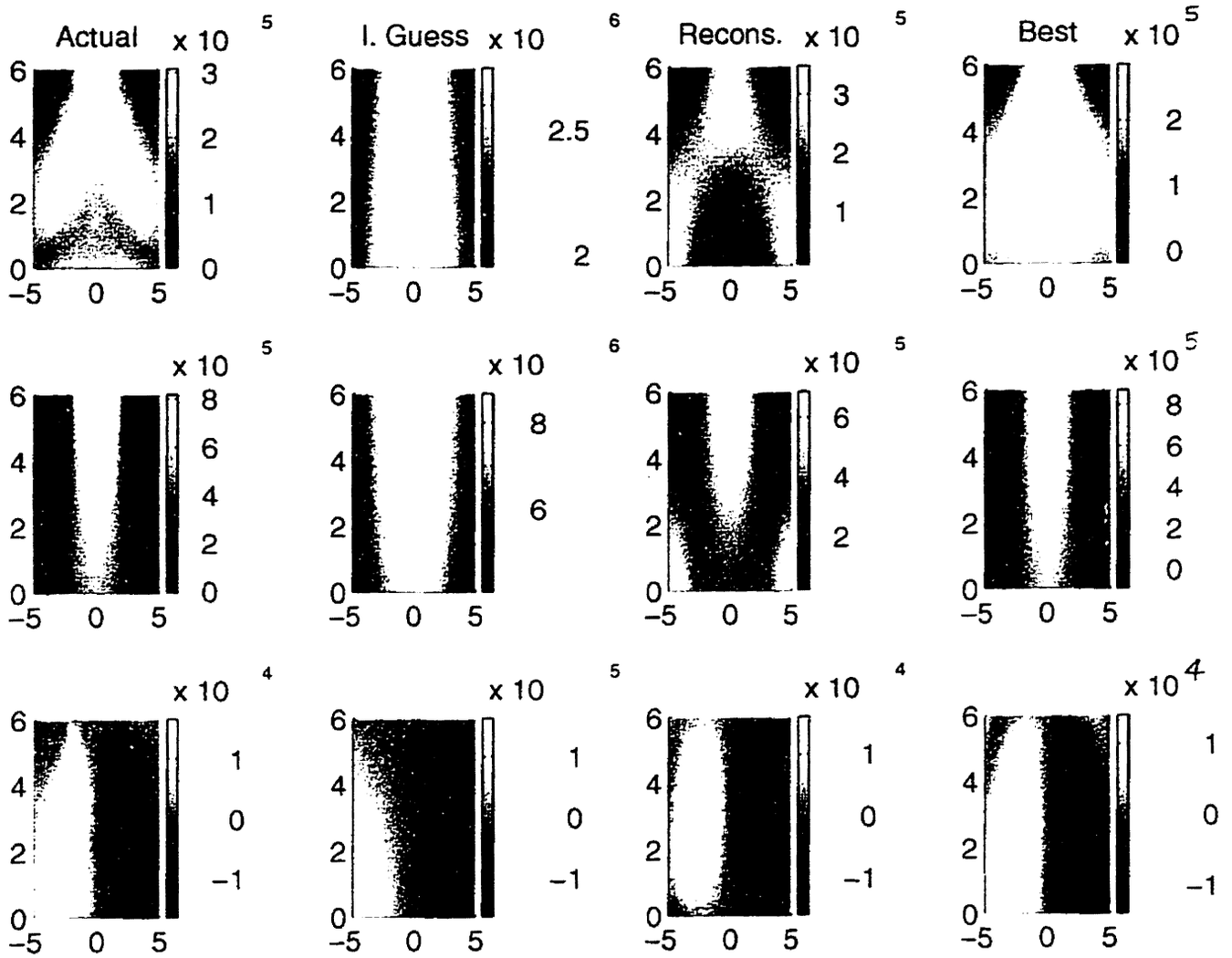


Figure 6.8: Distribution of σ_x, σ_y, and τ of the soil medium for the Actual, Initial, Reconstructed and the Best possible solution in the footing problem

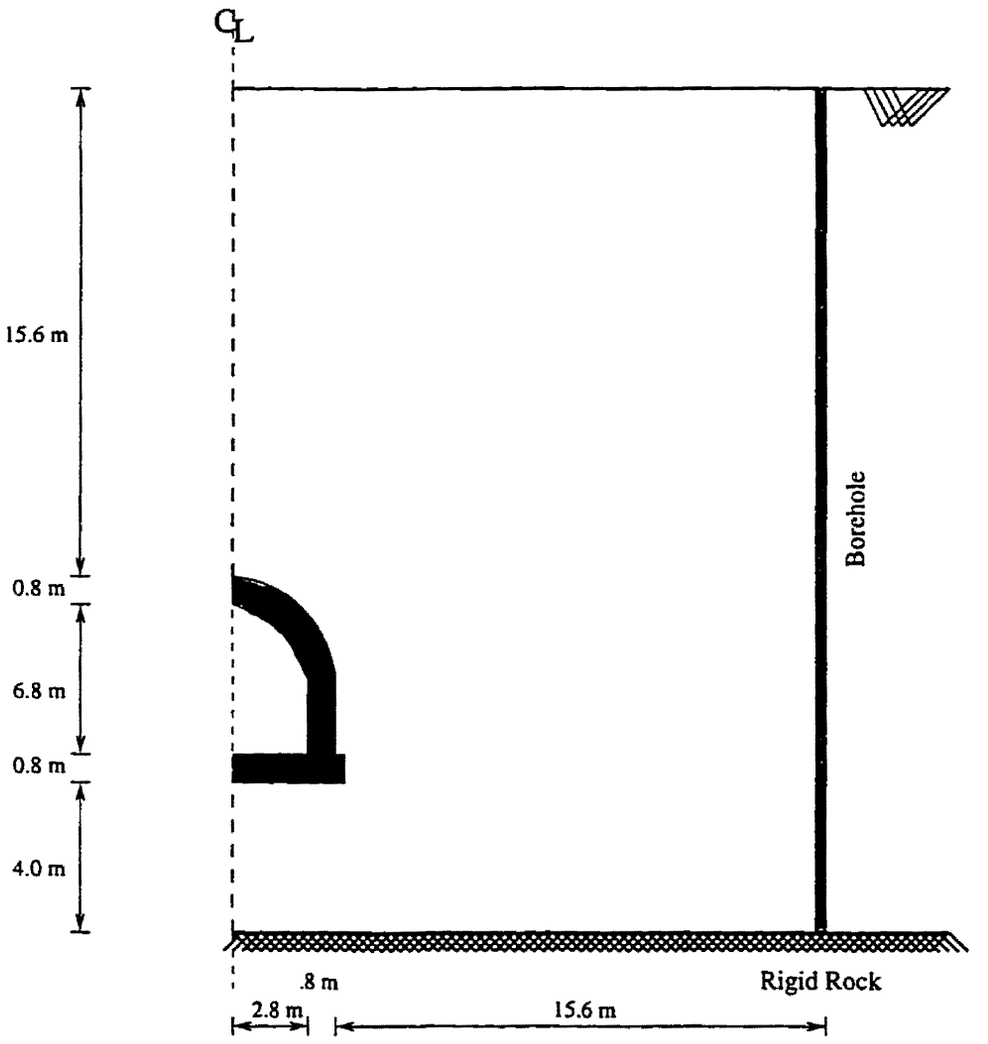


Figure 6.9: Soil region with a tunnel

$2.1 \times 10^2 \text{kg/cm}^2$ and Poisson's ratio $\nu = 0.3$. Yield is assumed to be governed by the Mohr-Coulomb surface, with a friction angle $\Phi = 30°$, and cohesion $C = 0.25\text{kg/cm}^2$. Using equations (6.1) to (6.3), we find the following Drucker-Prager criterion parameters corresponding to those of Mohr-Coulomb's.

$$\begin{aligned} \beta &= 39.76° \\ \sigma_C^\circ &= 0.5\text{kg/cm}^2 \end{aligned}$$

The soil around the tunnel is supported by a closed reinforced concrete shell as depicted in Figure 6.9. We assume that the concrete behaves as a linear elastic and isotropic medium with the following moduli:

$$\begin{aligned} E &= 2.86 \times 10^5 \text{kg/cm}^2 \\ \nu &= 0.15 \end{aligned}$$

By analyzing this problem using ABAQUS, we obtained the stress distributions of the soil medium as shown in the first column of Figure 6.10.

With the stress distribution results obtained from ABAQUS, we used the ray tracing code to determine travel-times. for this purpose, we assumed that $\alpha = 43.0$ and $\beta = 0.23$; we also assumed that two vertical boreholes are available on the boundaries of the region as depicted in Figure 6.9. Furthermore, it is assumed that 25 seismic wave sources are placed in one borehole, and the same number of receivers is placed in the other borehole. The sources and receivers are placed at $2.0m$ intervals beginning from the soil surface to $10.0m$ depth, and then they are placed at $1.0m$ intervals to $28.0m$ depth. Then we obtain 625 travel-times to be used for stress reconstruction in this problem.

Because of the symmetry in this problem, we used the same eighth order polynomial of equation (6.4) as the stress potential function.

Boundary Conditions

On the free surface of the soil region where $y = 0$, all the three stress components are equal to zero, i.e.

$$\begin{aligned} \sigma_x &= 0 \\ \sigma_y &= 0 \\ \tau &= 0 \end{aligned}$$

To incorporate these conditions and the linear nonlinear constraints required in the reconstruction algorithm, we define a set of discrete points on the region. In this problem, we used the nodal points of a grid of 0.4m intervals in both x and y directions.

Reconstructed Stress Field

In this example, the results are shown in Figure 6.10. The initial guess for stress distribution, shown in the second column of this figure, is obtained by making random perturbations, of amounts varying from %10 to %30, to the coefficients of the best possible solution.

Using the FORTRAN program we have developed based on the stress potential function approach, we obtained the reconstructed stress distributions shown in the third column of Figure 6.8 after seven iterations. Between the constructed stress distributions shown in the third column and the best possible solution shown in the fourth column, there is a very good agreement.

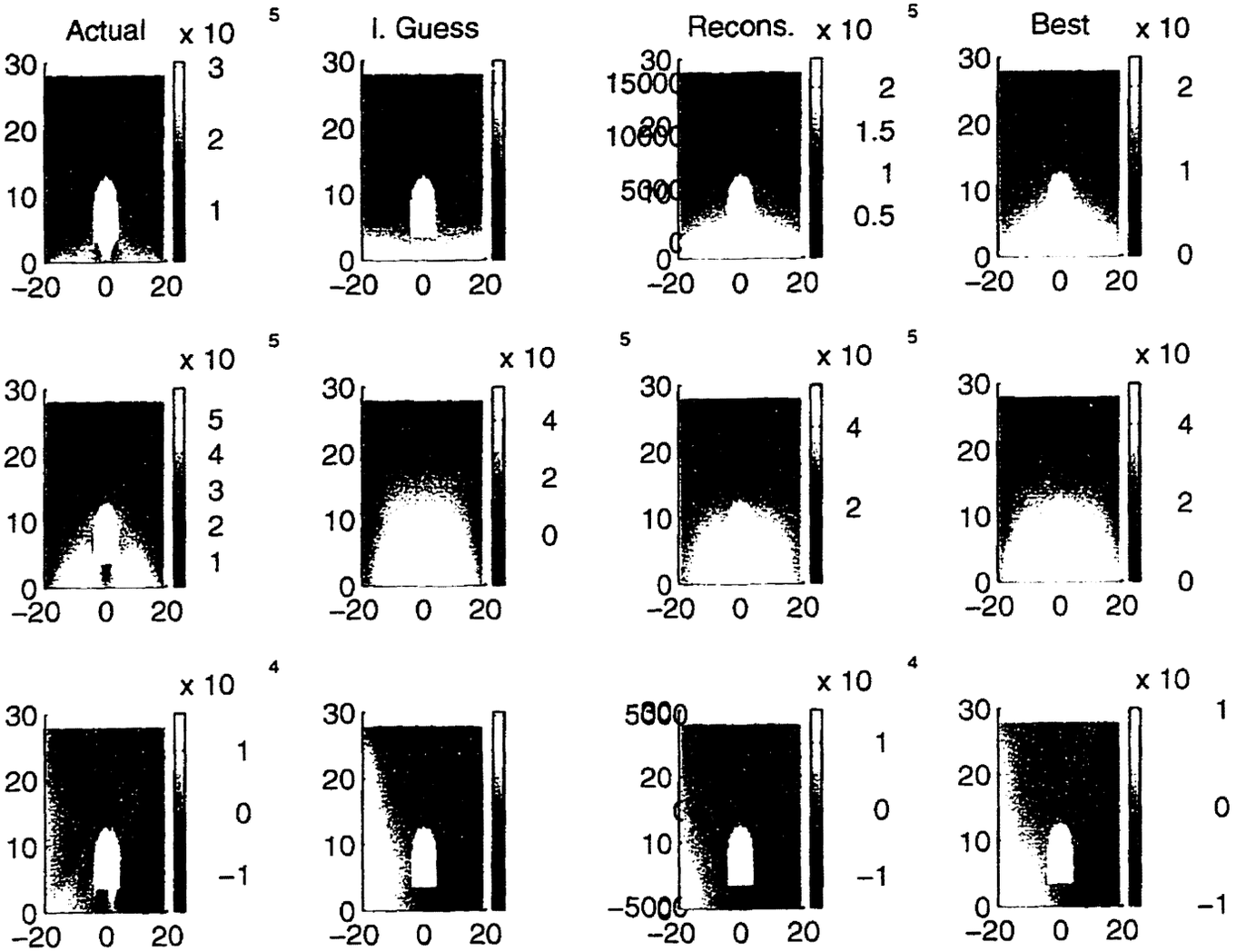


Figure 6.10: Distribution of σ_x, σ_y, and τ of the soil medium for the Actual, Initial, Reconstructed and the Best possible solution in the tunnel problem

6.3.4 Retaining Wall Problem Stress Reconstruction: Stress Functions Approach

In this example, as depicted in Figure 6.11, we have a soil medium supported by a retaining wall. Here, we are interested in stress reconstruction of the shaded area of the soil region. We assume that the soil region along with the retaining wall are long enough so that, for stress analysis, a plane strain model is satisfactory.

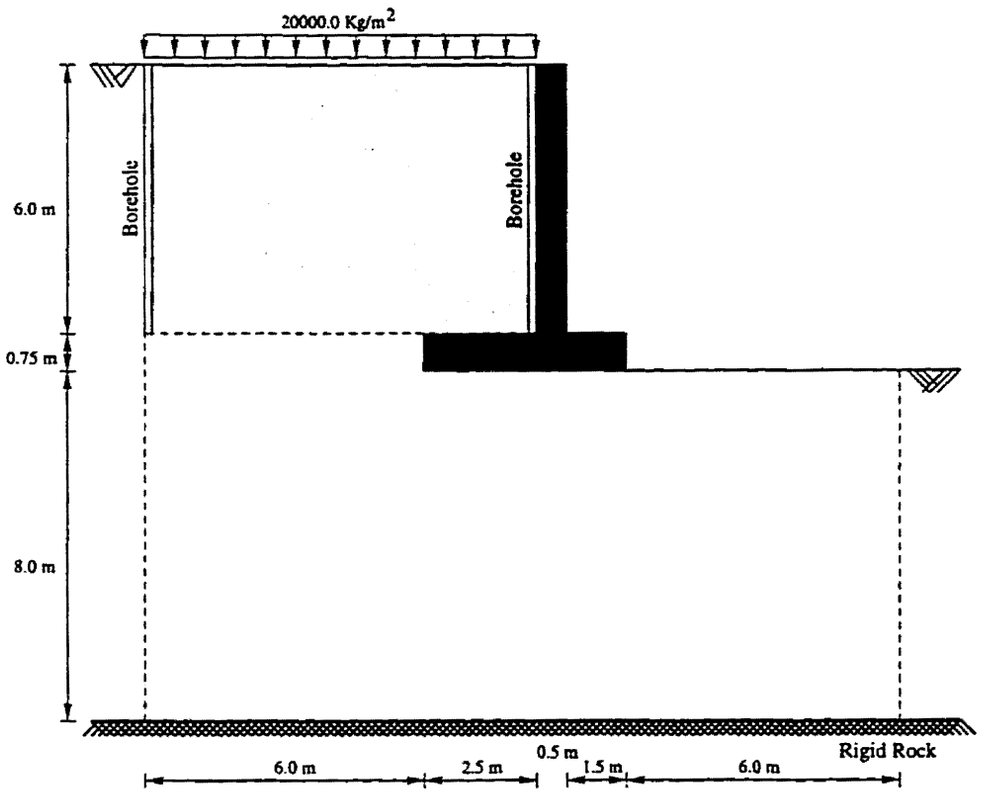


Figure 6.11: Soil region supported by a retaining wall

To obtain the stress distribution of the soil medium using ABAQUS, we assumed that soil with density $\gamma_s = 1.85\text{gr/cm}^3$ behaves as elastic-plastic material. The elastic response is assumed to be linear and isotropic with Young's modulus $E =$

$2.1 \times 10^2 \text{kg/cm}^2$ and Poisson's ratio $\nu = 0.3$. Yield is assumed to be governed by the Mohr-Coulomb surface, with a friction angle $\Phi = 30°$, and cohesion $C = 0.25\text{kg/cm}^2$. Using equations (6.1) to (6.3), we find the following Drucker-Prager criterion parameters corresponding to those of Mohr-Coulomb's.

$$\begin{aligned} \beta &= 39.76° \\ \sigma_C^\circ &= 0.5\text{kg/cm}^2 \end{aligned}$$

The soil region is supported by a reinforced concrete retaining wall. We assume that the concrete behaves as a linear elastic and isotropic medium with the following moduli:

$$\begin{aligned} E &= 2.86 \times 10^5 \text{kg/cm}^2 \\ \nu &= 0.15 \end{aligned}$$

By analyzing this problem using ABAQUS, we obtained the stress distribution of the soil medium. In the first column of Figure 6.12, the distribution of the stress components σ_x, σ_y and τ of the soil region of interest is shown.

With the stress distribution results obtained from ABAQUS, we used the ray tracing code to determine travel-times. for this purpose, we assumed that $\alpha = 41.0$ and $\beta = 0.20$; we also assumed that two vertical boreholes are available on the boundaries of the region as depicted in Figure 6.11. Furthermore, it is assumed that 9 seismic wave sources are placed in one borehole, and 10 receivers are placed in the other borehole. The sources and receivers are placed at $0.5m$ intervals beginning from $0.5m$ depth. Then we obtain 90 travel-times to be used for stress reconstruction in this problem.

As the problem is asymmetric, we use the following sixth order polynomial as the stress potential function:

$$\Phi = a_1x^2 + a_2xy + a_3y^2 + a_4x^3 + a_5x^2y$$

$$
\begin{aligned}
&+ a_6xy^2 + a_7y^3 + a_8x^4 + a_9x^3y + a_{10}x^2y^2 \\
&+ a_{11}xy^3 + a_{12}y^4 + a_{13}x^5 + a_{14}x^4y + a_{15}x^3y^2 \\
&+ a_{16}x^2y^3 + a_{17}xy^4 + a_{18}y^5 + a_{19}x^6 + a_{20}x^5y \\
&+ a_{21}x^4y^2 + a_{23}x^2y^4 + a_{24}xy^5 + a_{25}y^6
\end{aligned} \tag{6.8}
$$

This leads to the following relations for the stress field components:

$$
\begin{aligned}
\sigma_x &= \frac{\partial^2\Phi}{\partial y^2} + \gamma_s y \\
&= 2a_3 + 2a_6x + (6a_7 + \gamma_s)y + 2a_{10}x^2 \\
&\quad + 6a_{11}xy + 12a_{12}y^2 + 2a_{15}x^3 + 6a_{16}x^2y \\
&\quad + 12a_{17}xy^2 + 20a_{18}y^3 + 2a_{21}x^4 + 6a_{22}x^3y \\
&\quad + 12a_{23}x^2y^2 + 20a_{24}xy^3 + 30a_{25}y^4
\end{aligned} \tag{6.9}
$$

$$
\begin{aligned}
\sigma_y &= \frac{\partial^2\Phi}{\partial x^2} + \gamma_s y \\
&= 2a_1 + 6a_4x + (2a_5 + \gamma_s)y + 12a_8x^2 \\
&\quad + 6a_9xy + 2a10y^2 + 20a_{13}x^3 + 12a_{14}x^2y \\
&\quad + 6a_{15}xy^2 + 2a_{16}y^3 + 30a_{19}x^4 + 20a_{20}x^3y \\
&\quad + 12a_{21}x^2y^2 + 6a_{22}xy^3 + 2a_{23}y^4
\end{aligned} \tag{6.10}
$$

$$
\begin{aligned}
\tau &= \frac{\partial^2\Phi}{\partial x\partial y} \\
&= a_2 + 2a_5x + 2a_6y + 3a_9x^2 \\
&\quad + 4a_{10}xy + 3a_{11}y^2 + 4a_{14}x^3 + 6a_{15}x^2y \\
&\quad + 6a_{16}xy^2 + 4a_{17}y^3 + 5a_{20}x^4 + 8a_{21}x^3y \\
&\quad + 9a_{22}x^2y^2 + 8a_{23}xy^3 + 5a_{24}y^4
\end{aligned} \tag{6.11}
$$

Thus in this approach we need to determine only 25 unknowns to reconstruct the stress field. In case α and β are unknown, we must determine 27 unknowns.

Boundary Conditions

The surface of the soil region supported by the retaining wall is subjected to a vertical uniform load of 20000.0kg/m^2, therefore, at $y = 0$, the stresses σ_y nd τ are known and we have:

$$\begin{aligned} \sigma_y &= 20000\,\mathrm{kg/m^2} \\ \tau &= 0 \end{aligned}$$

To incorporate these conditions and the linear and nonlinear constraints required in the reconstruction algorithm, we define a set of discrete points on the region. In this problem, we used the nodal points of a grid of 0.25m intervals in both x and y directions.

Reconstructed Stress Field

In this example, the results are shown in Figure 6.12. The initial guess for stress distribution, shown in the second column of this figure, is obtained by making random perturbations, of amounts varying from %10 to %40, to the coefficients of the best possible solution.

Using the FORTRAN program we have developed based on the stress potential function approach, we obtained the reconstructed stress distributions shown in the third column of Figure 6.12 after five iterations. Between the constructed stress distributions shown in the third column and the best possible solution shown in the fourth column, there is a very good agreement.

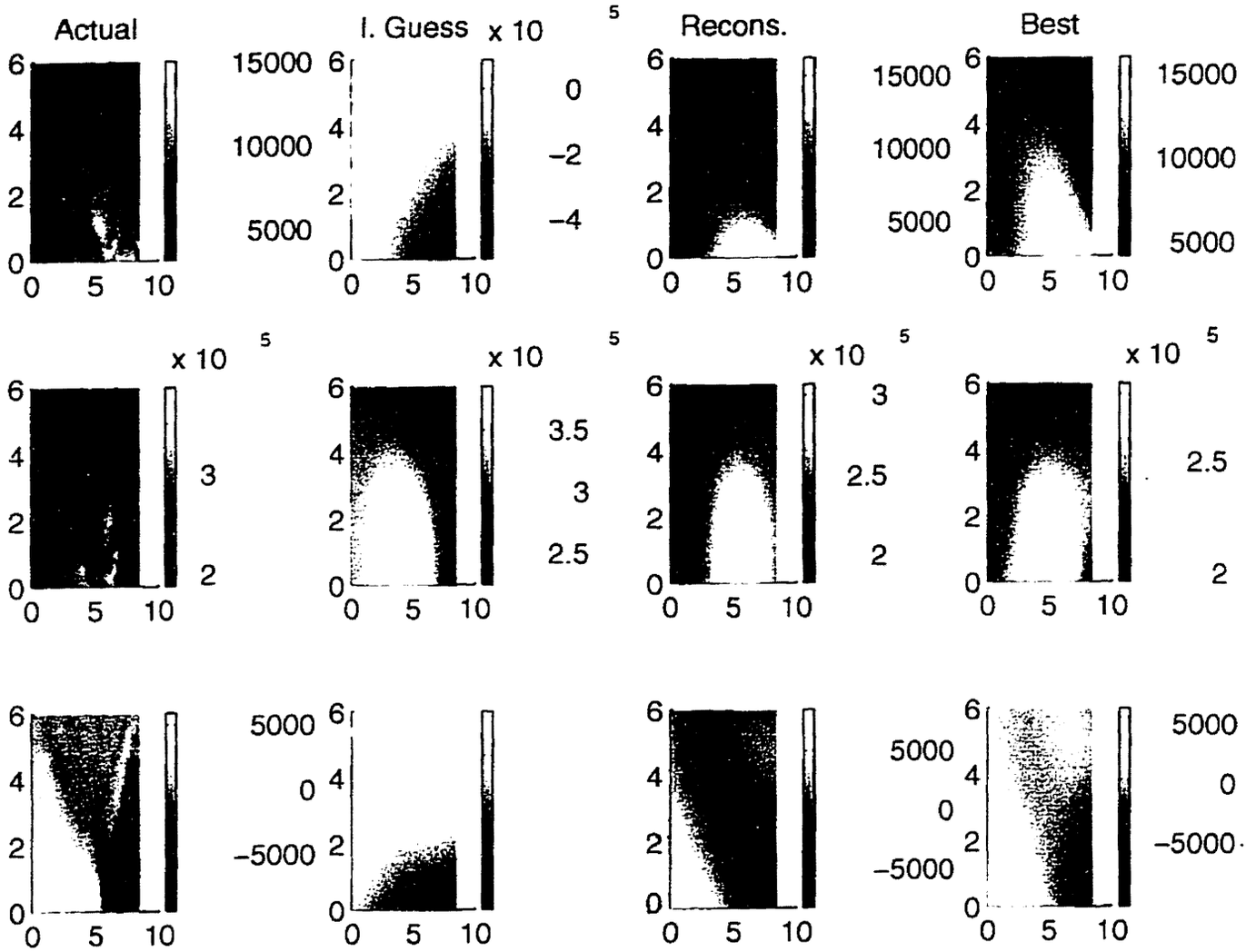


Figure 6.12: Distribution of σ_x, σ_y, and τ of the soil medium for the Actual, Initial, Reconstructed and the Best possible solution in the retaining wall problem

6.4 Closure

In this chapter, we have presented some results using the ray tracing and inversion techniques we have developed in this study.

In section 6.2, we presented the ray tracing results for a typical footing problem; these results are obtained using the ray tracing technique we have developed for stress induced anisotropic media. Since, to our knowledge, there is no other ray tracing technique available for stress induced anisotropic media, these results cannot be verified. However, the tendency of the ray paths to pass through the high stress zone is evident which somewhat confirms the performance of the ray tracing technique.

In section 6.3.1, we presented the results of applying the nodal stress approach for stress reconstruction in a small-scale problem. In that problem, the initial guess was obtained by perturbing the actual stress distribution by around %10. Under these circumstances, the reconstruction technique performance was acceptable. However, the approach failed for more distant starting guesses for this-small scale problem, and for any guess for large-scale ones.

In sections 6.3.2 to 6.3.4, we presented the results of stress reconstruction we obtained by applying the stress potential function approach in three typical geotechnical problems. These results indicate that the inversion technique based on this approach is reliable and capable of reconstructing general features of stress fields. However, due to some shortcomings in the polynomial representation of the stress field, delicate details in stress distribution, if any exist, cannot be reconstructed.

Chapter 7

Closure

This dissertation embodies the results of research which was motivated by the success of computerized tomography (CT) in various disciplines. In this research, we employed the concepts and techniques of CT to develop a model which can be used for monitoring stress distribution of granular soil media. This is an inverse problem, and we assume that travel-time data of compressional seismic waves is available for this purpose.

The media of interest in this problem can be categorized as inhomogeneous and anisotropic. Accordingly, we first derived a relation between travel-time of compressional waves and stress distribution. This relation is in the form of Fredholm equation of the first kind. As the second step, based on that equation, we developed a fast two-point ray tracing technique for solving the direct problem.

For the inverse problem, we first developed an inversion technique based on considering the unknown stress components on prespecified nodes on the region of interest. This technique led to a large scale highly nonlinear optimization problem which appeared to be impractical to handle real world problems. Thus, we

developed another inversion technique based on parametric representation of stress field. This technique led to a highly nonlinear, though manageable, optimization problem. Finally, based on synthetic data, we presented a number of illustrative examples to evaluate the performance of the models developed in this work.

It may be noted that the present work represents a part of that required to establish a more reliable and sophisticated model for soil stress imaging. Due to the novelty of this work, no simulated or experimental results are available in the literature to evaluate the performance of the models of this work; hopefully, this work will trigger further future research in this area.

Finally, for further research in this area, we suggest the following topics.

- Incorporating more realistic anisotropy models.
- Extending the presented models for fluid-saturated porous media.
- Encoding and evaluating the 3-D model presented in this work.
- Incorporating other methods of parametric representation of stress field in the stress potential function approach.
- Sometimes there is a known closed form solution to a problem which is similar to the one being studied. For example, there is a known closed form solution for a footing on an elastic half-space. In such a case, we may take this solution as a first approximation, and use a polynomial stress function to model the difference between our problem and the classic one.
- obtaining and using measured travel-time data for inversion.

Bibliography

[AK68] Z. Alterman and F. C. Karal. Propagation of elastic waves in layered media by finite differences methods. *Bulletin of Seis. Soc. Am.*, 58:367–398, 1968.

[AK82] A. H. Anderson and A. C. Kak. Digital ray tracing in two-dimensional refractive fields. *J. Acoust. Am.*, 72(5):1593–1606, 1982.

[AK84] A. H. Anderson and A. C. Kak. Simultaneous algebraic reconstruction technique (sart): A superior implementation of the art algorithm. *Ultrasond Imaging*, 6:81–94, 1984.

[AK93] E. Asakawa and T. Kawanaka. Seismic ray tracing using linear traveltime interpolation. *Geophysical Prospecting*, 41:99–111, 1993.

[Alo95] M. A. Aloufi. *Wave Propagation in Particulate Media - Micromechanics Experimental Study.* PhD thesis, University of Waterloo, Waterloo, Ontario, 1995.

[AR80] K. Aki and P. J. Richards. *Quantitative Seismology – theory and methods*, volume 1. W. H. Freeman and Company, San Francisco, 1980.

[AS95] M. A. Aloufi and J. C. Santamarina. Low and high strain macrobehaviour of grain masses - the effect of particle eccentricity. *ASAE*, 38(3), 1995.

[Aul73] B. A. Auld. *Acoustic Fields and Waves in Solids.* J. Wiley-Interscience Publication, New York, 2 edition, 1973.

[BBC+85] T. N. Bishop, K. P. Bube, R. T. Cutler, R. T. Langan, P. L. Love, J. R. Resnick, R. T. Shuey, D. A. Spindler, and H. W. Wyld. Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, 50:903–923, 1985.

[BBC89a] N. D. Bregman, R. C. Bailey, and C. H. Chapman. Crosshole seismic tomography. *Geophysics*, 54(2):200–215, 1989.

[BBC89b] N. D. Bregman, R. C. Bailey, and C. H. Chapman. Ghosts in tomography: The effects of poor angular coverage in 2-d seismic traveltime inversion. *Canadian Journal of Geophysics*, 97:7–27, 1989.

[BBC+94] R. Barrett, M. Berry, T. Chan, J. Demmel, J. Donato, J. Dongara, V. Eijkout, R. Pozo, C. Romine, and H. van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods.* SIAM, 1994.

[BC79] N. Bleinstein and J. K. Cohen. Direct inversion procedure for claerbout's equations. *Geophysics*, 44(6):1034–1040, 1979.

[BC90] B. S. Byun and D. Corrigan. Seismic traveltime inversion for transverse isotropy. *Geophysics*, 55(2):192–200, 1990.

[BCB89] N. D. Bregman, C. H. Chapman, and R. C. Bailey. Traveltime and amplitude analysis in seismic tomographycrosshole. *Journal of Geophysical Research*, 94(B6):7577–7587., 1989.

[BCG89] B. S. Byun, D. Corrigan, and J. E. Gaiser. Anisotropic velocity analysis for lithology discremention. *Geophysics*, 54:1564–1574, 1989.

[Ber89a] J. G. Berryman. Fermat's principle and nonlinear traveltime tomography. *Physical Review Letters*, 62(25):2953–2956, 1989.

[Ber89b] J. G. Berryman. Weighted least-squares criteria for seismic traveltime tomography. *IEEE Transactions on Remote Sensing*, 27(3):302–309, 1989.

[Ber90] J. G. Berryman. Stable iterative reconstruction algorithm for nonlinear traveltime tomography. *Inverse Problems*, 6:21–42, 1990.

[Ber92] J. G. Berryman. Seismic crosshole tomography and nonlinear constrained optimization. In *Geophysical Inversion, SIAM*, pages 396–414. SIAM, 1992.

[BG69] G. Backus and F. Gilbert. Constructing p-velocity models to fit restricted set of traveltime dtat. *Bull. Seism. Soc. Am.*, 59:1407–1414, 1969.

[BGL+87] R. P. Bording, A. Gersztenkorn, L. R. Lines, J. A. Scales, and S. Treitel. Applications of seismic traveltime tomography. *Geophys. J. R. astr. Soc.*, 90:285–303, 1987.

[Bri46] L. Brillouin. *Wave Propagation in Periodic Structures.* McGraw Hill, New York, 1946.

[Byu82] B. S. Byun. Seismic parameters for media with elliptical velocity dependencies. *Geophysics*, 47(12):1621–1626, 1982.

[Byu84] B. S. Byun. Seismic parameters for transversely isotropic media. *Geophysics*, 54:1908–1914, 1984.

[BW65] M. Born and E. Wolf. *Principles of Optics : Electromagnetic Theory of Propagation, Interference and Diffraction of Light.* Pergamon Press, Bath, U. K., 3 (revised) edition, 1965.

[Č72] V. Červený. Seismic rays and ray intensities in inhomogeneous anisotropic media. *Gephys. J. R. astr. Soc.*, 29:1–13, 1972.

[Cas96] G. Cascante. *Propagation of Mechanical Waves in Particulate Materials - Experimental Micromechanics.* PhD thesis, University of Waterloo, Waterloo, Ontario, 1996.

[CB79] J. K. Cohen and N. Bleinstein. Velocity inversion procedure for acoustic waves. *Geophysics*, 44(6):1077–108, 1979.

[CF84] V. Červený and P. Firbas. Numerical modeling and inversion of traveltimes of seismic body waves in inhomogeneous anisotropic media. *Gephys. J. R. astr. Soc.*, 76:41–51, 1984.

[Cha75] R. Chander. On tracing seismic rays with specified end points. *J. Geophysics*, 41:173–177, 1975.

[Che85] W. Chen. *Soil Plasticity : theory and implementation.* Elsevier, Amesterdam, 1985.

[CJ82] V. Červený and J. Jech. Linearized solutions of kinematic problems of seismic body waves in inhomogeneous slightly anisotropic media. *J. Geophys,*, 51:96–104, 1982.

[CMS91] C. S. Chang, A. Misra, and S. S. Sundaram. Properties of granular packings under low amplitude cyclic loading. *J. Soil Dynamics and Earthquake Engineering*, 10(4):201–211, 1991.

[CP92] C. H. Chapman and G. P. Pratt. Traveltime tomography in anisotropic media- i. theory. *Geophys, J. Int.*, 109:1–19, 1992.

[Dem87] J. W. Demmel. The smallest perturbation of a submatrix which lowers the rank and constrained total least squares problems. *SIAM J. Numer. Anal.*, 24(1):199–206, 1987.

[Dev82] A. J. Devaney. A filtered backpropagation algorithm for diffraction tomography. *Ultrason. Imaging*, 4:336–350, 1982.

[Dev84] A. J. Devaney. Geophysical diffraction tomography. *IEEE Trans. Geological Science, Special Issue on Remote Sensing*, GE-22:3–13, 1984.

[dH80] F. R. de Hoog. Review of fredholm equations of the first kind. In F. R. de Hoog and M. A. Lukas, editors, *The Application and Numerical Solution of Integral equations*, pages 119–134. Sijthoff and Noordhoff, Alphen aan den Rijn, 1980.

[DL79] K. A. Dines and R. J. Lytle. Computerized geophysical tomography. *Proceedings of the IEEE*, 67(7):1065–1073, 1979.

[Dol90] A. T. Dolovich. *Some New Prespectives in Photoelasticity and Curved-Ray Tomography.* PhD thesis, University of Waterloo, 1990.

[DS83] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, Inc., Englewood Cliffs, N. J., 1983.

[GK65] G. H. Golub and W. Kahan. Calculating the singular values and pseudoinverse of a matrix. *SIAM J. Numer. Anal. 2*, pages 205–224, 1965.

[GL81] A. George and J. W. H. Liu. *Computer Solution of Large Sparse Positive Definite Systems.* Prentice-Hall, Inc., Englewood Cliffs, N. J., 1981.

[GMS+88] E. P. Gill, W. Murray, M. A. Saunders, and M. H. Wright. Recent developements in constrained optimization. *Computational and Applied Mathematics*, 22:257–270, 1988.

[GMW81] P. Gill, W. Murray, and M. H. Wright. *Practical Optimization.* Academic Press, Inc., New York, 1981.

[God90] J. D. Goddard. Nonlinear elasticity and pressure-dependent wave speeds in granular media. In *The Royal Society of London, London, England*, number 430, pages 105–131, 1990.

[Gro93] C. W. Groetsch. *Inverse Problems in the Mathematical Sciences.* Friedr. Vieweg and Sohn Verlagsgesellschaft mblt, 1993.

[Gus78] F. G. Gustavson. Two fast algorithms for sparse matrices: Multiplication and premuted transposition. *ACM Transactions on Mathematical Softwares*, 4(3):250–269, 1978.

[GVL80] G. H. Golub and C. F. Van Loan. An analysis of the total least squares problem. *SIAM J. NUmer. Anal.*, 17(6):883–893, 1980.

[GVL89] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The John Hopkins University Press, 2 edition, 1989.

[GVT86] O. Gauthier, J. Virieux, and A. Tarantola. Two-dimensional nonlinear inversion of seismic waveforms : Numerical results. *Geophysics*, 51(7):1387–1403, 1986.

[Han92] P. C. Hansen. Analysis of discrete ill-posed problems by means of l-curve. *SIAM Rev.*, 34(4):561–580, 1992.

[Hel94] K. Helbig. *Foundations of Anisotropy for Exploration Seismics.* Pergamon Press, U. K., 1994.

[HO93] P. C. Hansen and D. P. O'leary. The use of l-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Comput.*, 14(6):1487–1503, 1993.

[HS95] Karlsson Hibbit and Inc. Sorenson. *ABAQUS, User's Manual*, 1995.

[Iva85] S. Ivanson. A study of methods for tomographic velocity estimation in the presence of low-velocity zone. *Geophysics*, 50:969–988, 1985.

[Iva87] S. Ivanson. Crosshole transmission tomography. In G. Nolet, editor, *SEISMIC TOMOGRAPHY with Applications in Global Seismology and Exploration Geophysics*, pages 159–188. D. Reidel Publishing Company, Dordrecht, 1987.

[JA68] B. R. Julian and D. L. Anderson. Traveltimes, apparent velocities and amplitudes of body waves. *Bull. Seism. Soc. Am.*, 58:339–366, 1968.

[JG77] B. R. Julian and D. Gubbins. Three-dimensional seismic ray tracing. *J. Geophys.*, 43:95–113, 1977.

[JGSJ75] S. A. Johanson, J. F. Greenleaf, F. A. Samayoa, W. A. Duck, and Spotrand J. Reconstruction of three-dimensional velocity fields and other parameters by acoustic ray tarcing. In *Ultrasonics Symposium (IEEE Cat. number = 75 CHO 944-450)*, pages 46–51, 1975.

[JK76] C. V. Jakowatz and A. C. Kak. Computerized tomographic imaging using x-rays and ultrasound. Research Report TR-EE 72-26, School of Electrical Engineering, Purdue University, 1976.

[Kac37] S. Kaczmarz. Angenaherte auflosung von systemen linearer gleichungen. *Bull. Acad. Pol. Sci. Lett. A*, 6-8A:355–357, 1937.

[Kak85] A. C. Kak. Tomographic imaging with diffracting and nondiffracting sources. In S. Haykins, editor, *Array Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1985.

[KS88] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, New York, 1988.

[KS91] Mal. A. K. and S. J. Singh. *Deformation of Elastic Solids.* Prentice-Hall, Inc., Englewood Cliff, N. J., 1991.

[KSK82a] D. P. Knox, II Stokoe, K. H., and S. E. Kopperman. Effects of state of stress on velocity of low-amplitude shear wave propagating along principal stress directions in dry sand. Report GR-82-23, Civil Engineering Department, University of Texas at Austin, Tex., 1982.

[KSK82b] S. E. Kopperman, II Stokoe, K. H., and D. P. Knox. Effects of state of stress on velocity of low-amplitude compression waves propagating along principal stress directions in dry sand. Report GR-82-22, Civil Engineering Department, University of Texas at Austin, Tex., 1982.

[Lan61] C. Lanczos. *Linear Differential Operators.* D. Van Nostrand, London, 1961.

[LD80] K. A. Lytle and R. J. Dines. Iterative ray tracing between boreholes for underground image reconstruction. *IEEE Transactions on Geoscince and Remote Sensing*, GE-18(3):234–240, 1980.

[LL77] D. L. Larger and R. J. Lytle. Determining a subsurface electromagnetic profile from high-frequency measurements by applying reconstruction technique algorithms. *Radio Sci.*, 12(2):249–260, 1977.

[LS91] Y. Luo and G. T. Schuster. Wave-equation traveltime inversion. *Geophysics*, 56(5):645–653, 1991.

[LVG95] L. P. Lebedev, I. I. Vorovich, and G. M. L. Gladwell. *FUNCTIONAL ANALYSIS with Applications in Mechanics and Inverse Problems.* Kluwer Academic Publishers, 1995.

[MA81] N. Mahdavi-Amiri. *Generalized Constrained Nonlinear Least Squares Problems: Algorithm Approach.* PhD thesis, The John Hopkins University, Baltimore, Maryland, 1981.

[MAB89] N. Mahdavi-Amiri and R. H. Bartels. Constrained nonlinear least squares: An exact penalty approach with projected structured quasi-newton updates. *ACM Transaction on Mathematical Softwares*, 15(3):220–242, 1989.

[Mar87] J. M. Martinez. An algorithm for solving sparse nonlinear least squares problems. *Computing*, (39):307–325, 1987.

[McM83] G. A. McMechan. Seismic tomography in boreholes. *Geophys. J. R. astr. Soc.*, 74:601–612, 1983.

[Men89] W. Menke. *Geophysical Data Analysis : Discrete Inverse Theory, Revised Edition.* Academic Press, California, 1989.

[MH91] R. J. Michelena and J. M. Harris. Tomographic traveltime inversion using natural pixels. *Geophysics*, 56(5):635–644, 1991.

[MMH93] R. J. Michelena, F. Muir, and J. M. Harris. Anisotropic traveltime tomography. *Geophysical Prospecting*, 4:381–412, 1993.

[Mor48] P. M. Morse. *Vibration and Sound.* McGraw-Hill, New York, 1948.

[Mor87] P. Mora. Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, 52(9):1211–1228, 1987.

[MV77] J. A. Meijerink and H. van der Vorst. An iterative solution method for linear systems of which the coefficients matrix ia a symmetric m-matrix. *Math. Comp.*, 31:148–162, 1977.

[NM65] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7:308–313, 1965.

[Nol85] G. Nolet. Solving or resolving inadequate and noisy tomographic systems. *Journal of Computational Physics*, 61:463–482, 1985.

[Nol87] G. Nolet. Seismic wave propagation and seismic tomography. In G. Nolet, editor, *SEISMIC TOMOGRAPHY with Applications in*

Global Seismology and Exploration Geophysics, pages 1–23. D. Reidel Publishing Company, Dordrecht, 1987.

[Off58] C. B. Officer. *Introduction to the Theory of Sound Transmission.* McGraw-Hill, New York, 1958.

[PBJ+88] W. R. Petrick, D. T. Borup, S. A. Johnson, and M. J. Berggren. Seismic borehole tomography using full waveform inversion. In *58 th Ann. Internat. Mtg. Soc. Expl. Geophys., Expanded Abstracts*, pages 1250–1252, 1988.

[PC92] G. P. Pratt and C. H. Chapman. Traveltime tomography in anisotropic media- ii application. *Geophys, J. Int.*, 109:20–37, 1992.

[PD87] E. Petrakis and R. Dorby. Micromechanical modeling of granular soil at small strain by array of elastic spheres. Report CE-87-02, Department of Civil Engineering, Rensselaer Polytechnic Institute, Tory, NY, 1987.

[PF91] W. S. Phillips and M. C. Fehler. Traveltime tomography : A comparison of popular methods. *Geophysics*, 56(10):1639–1649, 1991.

[PG91] R. G. Pratt and N. R. Goutly. Combining wave-equation imaging with traveltime tomography to form high-resolution images from crosshole data. *Geophysics*, 56(2):208–224, 1991.

[PLK80] V. Pereyra, W. H. xK. Lee, and H. B. Keller. Solving two-point seismic-ray tracing problems in a heterogeneous medium. *Bulletin of the Seismological Society of America*, 70(4):79–99, 1980.

[Pra90] R. G. Pratt. Inverse theory applied to multi-source cross-hole tomography. part 2 : Elastic wave-equation method. *Geophysical Prospecting*, 38:311–329, 1990.

[PS82] C. C. Paige and M. A. Saunders. Lsqr: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Softwares*, 8(1):43–71, 1982.

[PTE88] W. A. Prothero, W. J. Taylor, and J. A. Eickemeyer. A fast two-point three-dimensional ray tracing algorithm using a simple step search method. *Bulletin of the Seismological Society of America*, 78(3):1190–1198, 1988.

[PTVF92] W. H. Press, S. a. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in FORTRAN; The Art of Scientific Computing.* Cambridge University Press, 2 edition, 1992.

[PW88] R. G. Pratt and M. H. Worthongton. The application of diffraction tomography to cross-hole seismic data. *Geophysics*, 53(10):1284–1294, 1988.

[PW90] R. G. Pratt and M. H. Worthington. Inverse theory applied to multi-source cross-hole tomography. part 1 : Acoustic wave-equation method. *Geophysical Prospecting*, 38:311–329, 1990.

[RB89] L. Rothenburg and R. J. Bathurst. Analytical study of induced anisotropy in idealized granular materials. *Géotechnique*, 39(4):601–614, 1989.

[RB92] L. Rothenburg and R. J. Bathurst. Micromechanical features of granular assemblies with planar elliptical particles. *Géotechnique*, 4(1):79–95, 1992.

[Roe79] S. K. Roesler. Anisotropic shear modulus due to stress anisotropy. *ASCE, Geotechnical Engineering Division, 105, GT7*, pages 871–880, 1979.

[Rot80] L. Rothenburg. *Micromechanics of Idealized Granular System.* PhD thesis, Carlton University, Ottawa, 1980.

[SC96] J. C. Santamarina and G. Cascante. Stress anisotropy and wave propagation: A micromechanical view. *Canadian Geotechnical Journal*, (33):770–782, 1996.

[Sch78] H. Schmertmann, J. Effect of shear stress on the dynamic bulk modulus of sand. Technical Report S-78-16, Waterways Experimental Station, 1978.

[Sch88] K. Schittkowski. Solving constrained nonlinear least squares problems by a general purpose sqp-method. In K. H. Hoffmann, J.-B. Hiriart-Urruty, C. Lemarechal, and J. Zowe, editors, *Trends in Mathematical Optimization.* Birkhauser Verlag Basel, 1988.

[SF95] J. C. Santamarina and M. Fam. Changes in dielectric permittivity and shear wave velocity during concentration diffusion. *Canadian Geotechnical Journal*, (32):647–659, 1995.

[SGM+93] J. C. Santamarina, J. Graham, C. Mac Dougall, and V. Roy. Tomographic imaging changes in effective stress in granular media (simulation study). *Transportation Research Record*, (1415):95–99, 1993.

[SP94] J. C. Santamarina and B. Potts. On the imaging of stress changes in particulate media: An experimental study. *Canadian Geotechnical Journal*, (31):215–222, 1994.

[Sta72] O. N. Stavroudis. *The Optics of Rays, Wavefronts, and Caustics.* Academic, New York, 1972.

[Tan71] K. Tanabe. Projection method for solving a singular system. *Numer. Math.*, 17:203–214, 1971.

[Tan87] T. Tanimoto. Surface - wave ray tracing equation and fermat's principle in an anisotropic earth. *Geophys, J. R. astr. Soc.*, (88):231–24, 1987.

[Tar84a] A. Tarantola. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266, 1984.

[Tar84b] A. Tarantola. The seismic reflection inverse problem. In F. Santosa, Y. H. Pao, W. Symes, and C. Holland, editors, *Inverse Problem of Acoustic and Elastic Waves*, pages 104–181. SIAM, 1984.

[TE80] C. H. Thurber and W. L. Ellsworth. Rapid solution of ray tracing problem in heterogeneous media. *Bulletin of the Seismological Society of America*, 70(4):1137–1148, 1980.

[TG82] S. P. Timoshenko and J. N. Goodier. *Theory of Elasticity.* McGraw-Hill, 3 edition, 1982.

[UT87] J. Um and C. H. Thurber. A fast algorithm for two-point ray tracing problem in heterogeneous media. *Bulletin of the Seismological Society of America*, 77(3):972–986, 1987.

[VAK92] A. Vafidis, F. Abramovici, and E. R. Kanasewich. Elastic wave propagation using fully vectorized high order finite-difference algorithms. *Geophysics*, 57(2):218–232, 1992.

[VHV88] S. Van Huffel and J. Vandewelle. The partial total least squares algorithm. *Journal of Computational and Applied Mathematics*, 21:333–341, 1988.

[Vid88] J. Vidale. Finite-difference calculation of traveltimes. *Bulletin of the Seismological Society of America*, 78:2062–2076, 1988.

[Vir84] J. Virieux. Sh-wave propagation in heterogeneous media, velocity-stress finite-difference method. *Geophysics*, 49:1933–1957, 1984.

[Vir86] J. Virieux. Psv-wave propagation in heterogeneous media, velocity-stress finite-difference method. *Geophysics*, 51:889–901, 1986.

[Wes71] R. L. Wesson. Traveltime inversion for laterally inhomogeneous crustal velocity models. *Bull. Seism. Soc. Am.*, 61, 1971.

[Whi83] J. E. White. *UNDERGROUND SOUND - Application of Seismic Waves.* Elsevier Science Publishers B. V., Amsterdam, The Netherlands, 1983.

[Whi89] D. J. White. Two-dimensional seismic refraction tomography. *Geophysical Journal*, 97:223–245, 1989.

[Wie87] E. Wielandt. On the validity of the ray approximation for interpreting delay times. In G. Nolet, editor, *SEISMIC TOMOGRAPHY with Applications in Global Seismology and Exploration Geophysics*, pages 85–98. D. Reidel Publishing Company, Dordrecht, 1987.

[Wig72] R. A. Wiggins. The general linear inverse problem : Implication of surface waves and free oscillations for earth structure. *Rev. Geophys. and Space Phys.*, 10:251–285, 1972.

[WM88] A. Witten and J. E. Molyneux. Geophysical imaging with arbitrary source illumination. *IEEE Trans. on Geoscince and Remote Sensing*, 26(4):409–419, 1988.

[WR88] M. J. Woodward and F. Rocca. Wave-equation tomography. In *58 th Ann. Internat. Mtg. Soc. Expl. Geophys., Expanded Abstracts*, pages 1232–1235, 1988.

[YL76] J. P. Yang and W. H. K. Lee. Preliminary investigations on computational methods for solving two points seismic ray tracing problems in a heterogeneous and isotropic medium. *U. S. Geol. Surv., Open-File Rept. 76*, 1976.

[YR84] P. Yu and F. E. Richart. Stress ratio effects on shear modulus of dry sands. *Geotechnical Engineering, ASCE*, 110(3):331–345, 1984.