INVERSE METHODS IN SEISMOLOGY

by

JEAN SCHEPERS DE VILLIERS

submitted in accordance with the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in the subject

PHYSICS

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF M BRAUN

CO-SUPERVISOR: DR P POTGIETER

NOVEMBER 2009

Abstract

The problem of fitting a material property of the earth to a certain model by analysing a returned seismic signal is investigated here. Analysis proceeds with methods taken from the theory of inverse problems. Seismic wave inversion is tackled by minimisation of the objective function with respect to the model parameters. Absorbing boundary conditions are implemented using an exponentially decaying ansatz.

DECLARATION

I declare that *Inverse Methods in Seismology* is my own work and that all sources that I have used or quoted have been indicated and acknowledged by means of references.

Signature _____ Date _____

FINANCIAL ASSISTANCE

The financial assistance of the Department of Labour (DoL) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the DoL.

ACKNOWLEDGEMENTS

Firstly, the glory goes to THE LORD in heaven who gave me the strength, perseverence and courage to study the physical universe He has created.

My utmost appreciation goes to `my parents, brother` and `other family members`, both near and from afar, for giving me the moral support without which this thesis wouldn't have seen the light of day.

My sincere gratitude to my supervisor Prof. M. Braun for introducing me to geophysics and especially the seismology field. My deepest appreciation goes to him for his guidance, advice and patience during this time. I am also grateful to him for assistance in matters of computation, the contract and other administrative duties. His moral support throughout the course of this project is much appreciated.

My thanks also goes to my co-promoter Dr. Paul Potgieter at the UNISA Decision Sciences department who has shown an interest in what I was doing.

Thanks to the **National Research Foundation** for awarding me the Department of Labour (DoL) Scarce Skills Scolarship for students with disabilities, and the **UNISA Financial Aid Bureau** for the supplementary bursary respectively in order for me to further my studies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Seismology is the study of wave propagation through the earth (right down to its core) and this can include earthquakes, volcanoes and the tectonic plates in particular. Seismic studies have also been done based on man-made activities, such as detonations or prospecting for metals, minerals and oil. An introductory treatment into this field and elasticity theory is given in sections 2.1 and 2.2, and more extensively dealt with in [1, 2] as well as [3, 4, 5].

One easy and direct way of approaching the topic is to predict what the observed signal will look like, given an already known and surveyed slab of the earth's crust. One could set up a simplified model mimicking the referred to region, and calculate the recorded waves, much like in quantum scattering theory. But more often than not, there are regions and slabs of the earth's crust, especially the ocean's bottom, that still remain unexplored for mining purposes. What is much more desirable is the reverse process: deriving whatever is going on inside the crust, given the recorded signals. Which is where inverse methods would come in. An introductory text on inverse problems as a mathematical theory is given by Kirsch [6] This inversion theory is more indirect and can only be achieved numerically in most situations.

Inverse problems have a long history in accoustics, geophysics, optics and electromagnetics. Some applications of the theory are described next. Wand [7] uses a Bayesian statistical inference method on stochastic inverse problems in thermo-

dynamics, especially heat transfer. Wei [8] introduces a filtering inverse method for their unequal travel-time layered model to filter out any unwanted or spurious reflections. The book of Taroudakis and Makrakis [9] contains a colection of contemporary applications for the theory in underwater accoustic porblems, since signals from ocean sensors have become more numerous and sophisticated enough for the realistic identification of the ocean's parameters. Aster [10] focusses on a frequentist statistical inverse approach, discussing such statistical tests as the Student-t and Chi-square, and compares the theory to the Bayesian methods. Bayesian inverse theory is based on a probabilistic approach, but suffers from a drawback in that an *ad-hoc* prior distribution for the model is chosen from intuitive experience and that is considered in the theory as uninformative. Tarantola [11] gets around this problem by choosing a homogeneous probability distribution and posits that it is as informative as any other prior distribution.

The aim of this study is to use minimisation techniques, based on the inverse problem theory, for the seismic inversion of the material properties (MP) to within $\pm 1\%$ of the difference between material properties at the surface and the substratum. These properties are prescribed by model input parameters $\mathbf{m}$ through the Fourier method or finite element method (FEM) as a material property (MP) function (see below). An optimised amplitude is constructed from $\mathbf{m}$ to fit a target function. The optimised amplitude is constructed through the differential wave equation containing the MP function, and the ODE represented succintly by the operator $\mathbf{G}(\mathbf{m})$. The target amplitude, represented by a (synthetic) data function $\mathbf{d}$ instead, is constructed the same way but with a MP function already given analytically in advance. The least-square objective function is then a sum of squared residuals, $\mathbf{r}^2 = (\mathbf{G}(\mathbf{m}) - \mathbf{d})^2$, and must be minimised by any least-squares regression technique. (See [10, 11] for the notation, and section 3.5 for more details).

With this goal in mind, it is possible that, as with quantum inverse scattering problems, one could reconstruct a profile of the composition of the earth's crust from given seismic observations, then apply a model of its composition as a best

fit to the data, using inversion (minimisation) techniques. Much observation on or below the ocean bottom have brought out the need for a more theoretical basis to explain these data. In this work a better technique is sought to predict seismic signals reflected from the sea floor in these experiments. Our purpose would be to develop a technique that can approach any given complex physical setup. To achieve this aim we first concentrate on simple physical systems, then work our way towards the more complex one's. All along the way testing them with the actual observations.

This thesis follows on where B.R. Mabuza left off in [12], but now moves in a different direction. Where Mabuza used quantum inverse scattering theory to tackle the macroscopic seismic inversion problem, we are using minimisation techniques, such as those described in [13, 14], and implemented with the Merlin package [15, 16]. This is to achieve inversion of a similar seismic problem modelled after the inverse problem theory with a probabilistic approach first started in [17] and developed to the full in [11] by Tarantola. This is used to estimate material properties in a medium. The minimisation techniques used to invert non-linear models setup are, for example, the following: The quasi-newton methods [18, 19], conjugate gradient method [20], geometric methods (i.e. simplex) [21], and Levenberg-Marquardt method [22]. This last technique is best suited for the sum-of-squares objective function in the minimization.

The finite difference method (FDM) [23, 24, 25] is used for the solution of the wave equation. Its real use comes when it is applied to functions that can only be calculated numerically. Analytic expressions can be differentiated analytically and the result of the derivative simply given as new expressions. Numerically derived functions must be differentiated with the FDM. There are three varieties: forward (0,+h), central (-h,+h) and backward (-h,0). We apply mostly the central flavour as will be explained in section 3.1.

The finite element method (FEM) can be applied to any difficult problem where a function has part or all of its domain cut up into elements. Smaller elements are

concentrated at regions of higher slopes of the function for easier approximation, and vice versa for larger elements. The function is approximated by basis functions of a chosen order in each element. This method was actually invented by structural engineers for complex elasticity projects with plate and shell elements [26, 27, 28]. Non-linear material behaviour was also modelled and not only static but dynamic problems were tackled at the time. Apparently unaware of this, the applied mathematicians did similar studies, but using variational principles. Once a convergence proof was given, it was realised the method could be interpreted in terms of variational techniques and the two communities became aware of each other. The method became important, thereafter, not only from a practical but also a theoretical point of view. Academics from the natural sciences then applied it to problems such as weighted residuals, viscous fluid flow, electromagnetic fields and the coupled diffusion-convection problem involved in solving the Navier-Stokes equations. Here we introduce the finite element method only in three dimensions.

However Falk [24] and Kosloff [29] pointed to some drawbacks of the FDM and FEM. Falk [24] used the FDM on fluid-filled borehole modelling schemes, where grid spacing was coarse with respect to the accurate representation of the shortest wavelength. Difficulties originating from the large-scale differences between grid spacing and the size of their borehole were overcome by a grid refinement technique. Kosloff [29] used a fast Fourier Transform method (FFTM) to calculate spatial derivatives instead of finite differences. The resulting operators were more accurate, requiring only two grid points to resolve the spatial wavelength.

The spectral element method (SEM) was introduced by Patera [30] for fluid dynamics and applied to elastic wave equations in realistic applications to two-dimension and three-dimension seismic problems in [31]. It combined the generality of the FEM, using a high order polynomial basis, with good convergence properties of the spectral method.

We introduce the finite element method only in three dimensions for compar-

ison. We used the Fourier series approximation to optimise the material property function in one and two dimensions only. Such a function should be smooth, no discontinuities allowed, and can only work on a small region around the space point on which it is applied. The Fourier coefficients are the MERLIN input parameters to be optimised.

In order that boundary reflections do not interfere with those of the wave within the body of a numerical grid, these reflections must be eliminated. Originally this was attempted by simply adding an *ad hoc* damping function term to the wave equation [32]. That did not succeed in eliminating the reflection completely. Kleefsman [33], in her investigation of the stormy sea's impact on research ships, applied a numerical beach region with the addition of a pressure dissipation term. Vacus [34] studied the corner problem of two intersecting boundaries with absorbing conditions. Hagstrom used Radiation boundary conditions in his wave simulations [35]. Further absorbing boundary conditions are described by Givoli in [36]. However, in this study a new approach is implemented for a non-reflecting boundary condition [37]. This is based on Berenger's new NRBC called the perfectly matched layer (PML) [38] and already applied in [39] by Michou.

This thesis proceeds as follows. In chapter 2 the theory of elasticity and the differential equations that govern wave propagation are discussed. Chapter 3 describes firstly the numerical methods to assist in the differentiating, approximating or optimising the functions involved. Secondly, a method for absorbing the wave at the boundaries is introduced. Finally, and primarily, the techniques, especially Levenberg-Marquardt, that recover material properties via minimisation are described. The results for the material properties are summarised in chapter 4 where the methods are tested. The Fourier series is implemented only in one and two dimensions with parallel layers. Only the finite element method is used in three dimensions with a feature. Finally in chapter 5 we give conclusions based upon these results.

Chapter 2

Theoretical Framework

Here we discuss the elasticity theory, where we begin with the symmetric stress as a surface traction and strain in deformations. These tensors are used in Hooke's law with the Lamé parameters, and the material properties can be derived from certain experimental setups. Finally, we separate the wave equation via using Helmholtz potentials into two partial differential equations representing longitudinal and transverse waves. To simplify matters, we only consider propagation through fluids and gas where the first Lamé parameter is non-homogeneous, the second one is zero, and the vector potential transverse wave equation falls away.

2.1 Stress

The following is inspired by the treatment in [3]. Let V be a region bounded by its surface S, occupied at time t by a material body. $\mathbf{n}$ is the outward normal unit vector of S, and there exist two vector fields. One is called the body force $\mathbf{f}(\mathbf{x})$ over V. The other is called the stress vector $\mathbf{t}(\mathbf{x}, \mathbf{n})$ over S, also known as surface traction. Then the total force on the volume region is

$$\int_S \mathbf{t}(\mathbf{x}, \mathbf{n}) dS + \int_V \rho \mathbf{f}(\mathbf{x}) dV, \tag{2.1}$$

where ρ is the mass density at point $\mathbf{x}$.

Next we define the linear momentum density to be $\rho\mathbf{v}(\mathbf{x})$ per unit volume, where $\mathbf{v}(\mathbf{x})$ is the velocity of the body and ρ is assumed to be constant for the

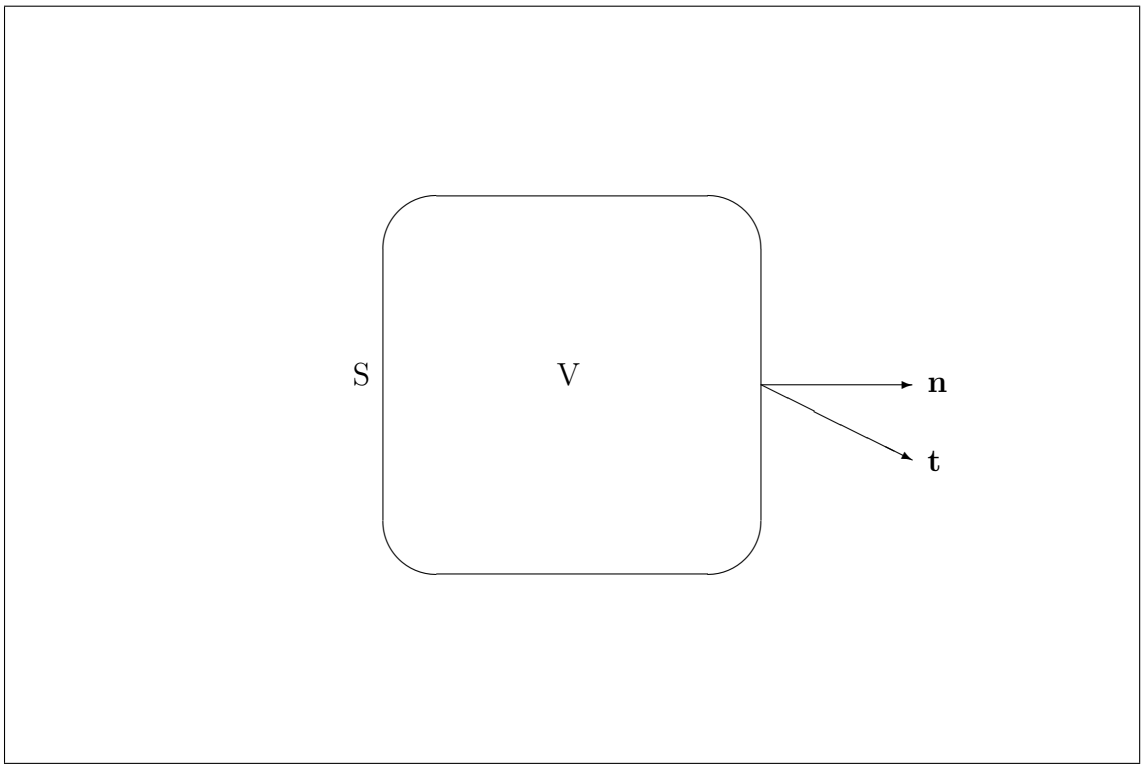

Figure 2.1: Region V bounded by surface S with normal $\mathbf{n}$ and traction $\mathbf{t}$.

region considered, so that the total linear momentum of the body is

$$\rho \int_V \mathbf{v}(\mathbf{x}) dV. \tag{2.2}$$

The rate of change of the linear momentum (Eq.2.2) is taken to be equal to the resultant total force (Eq.2.1) on the body, i.e.

$$\int_V \rho \frac{d\mathbf{v}}{dt} dV = \int_S \mathbf{t}(\mathbf{x}, \mathbf{n}) dS + \int_V \rho \mathbf{f}(\mathbf{x}) dV. \tag{2.3}$$

In this work we assume that the medium is at rest.

Suppose now that region V is cut into two parts V_1 and V_2 by a surface s. Region V_1 is bounded by s and a part S_1 of S, and similarly for V_2. As the relationship holds for any part of the body, it must hold for regions V_1 and V_2 separately. $\mathbf{m}$ denotes the unit normal to s out of V_1 into V_2. Therefore

$$\int_{V_1} \rho \frac{d\mathbf{v}}{dt} dV = \int_s \mathbf{t}(\mathbf{x}, \mathbf{m}) dS + \int_{S_1} \mathbf{t}(\mathbf{x}, \mathbf{n}) dS + \int_{V_1} \rho \mathbf{f}(\mathbf{x}) dV$$

and

$$\int_{V_2} \rho \frac{d\mathbf{v}}{dt} dV = \int_s \mathbf{t}(\mathbf{x}, -\mathbf{m}) dS + \int_{S_2} \mathbf{t}(\mathbf{x}, \mathbf{n}) dS + \int_{V_2} \rho \mathbf{f}(\mathbf{x}) dV.$$

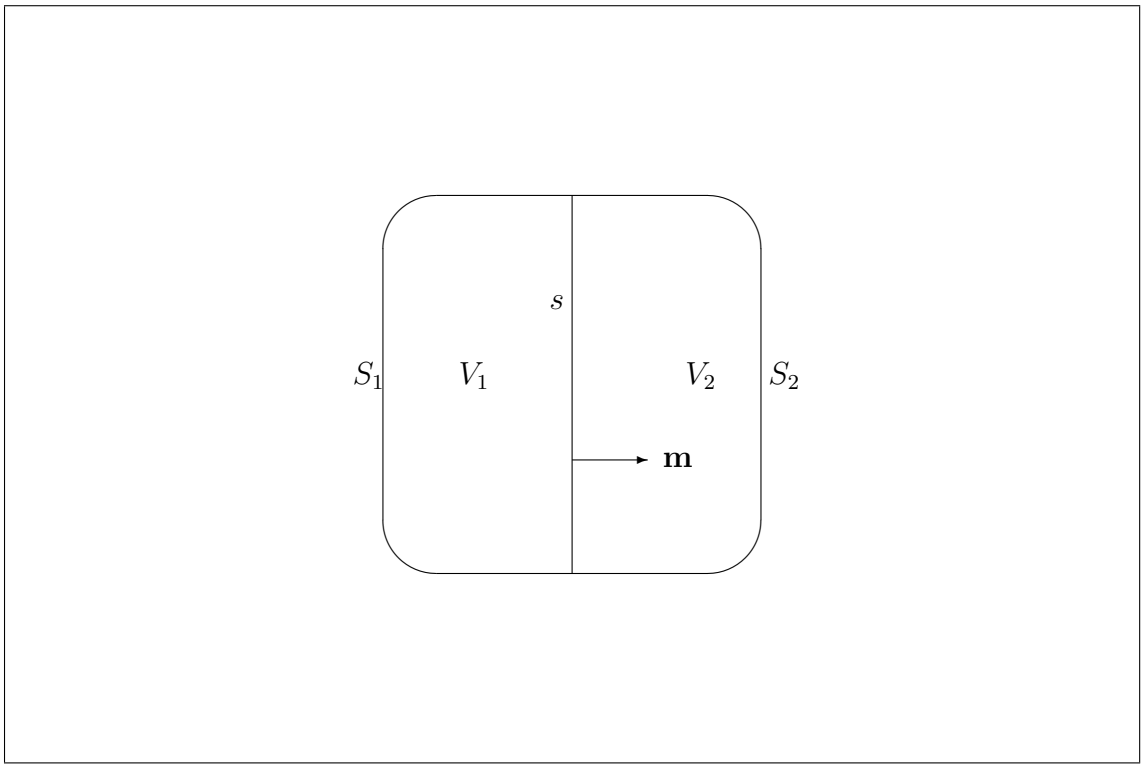

Figure 2.2: Regions V_1 and V_2 with common surface s and normal $\mathbf{m}$ and outer surfaces S_1 and S_2.

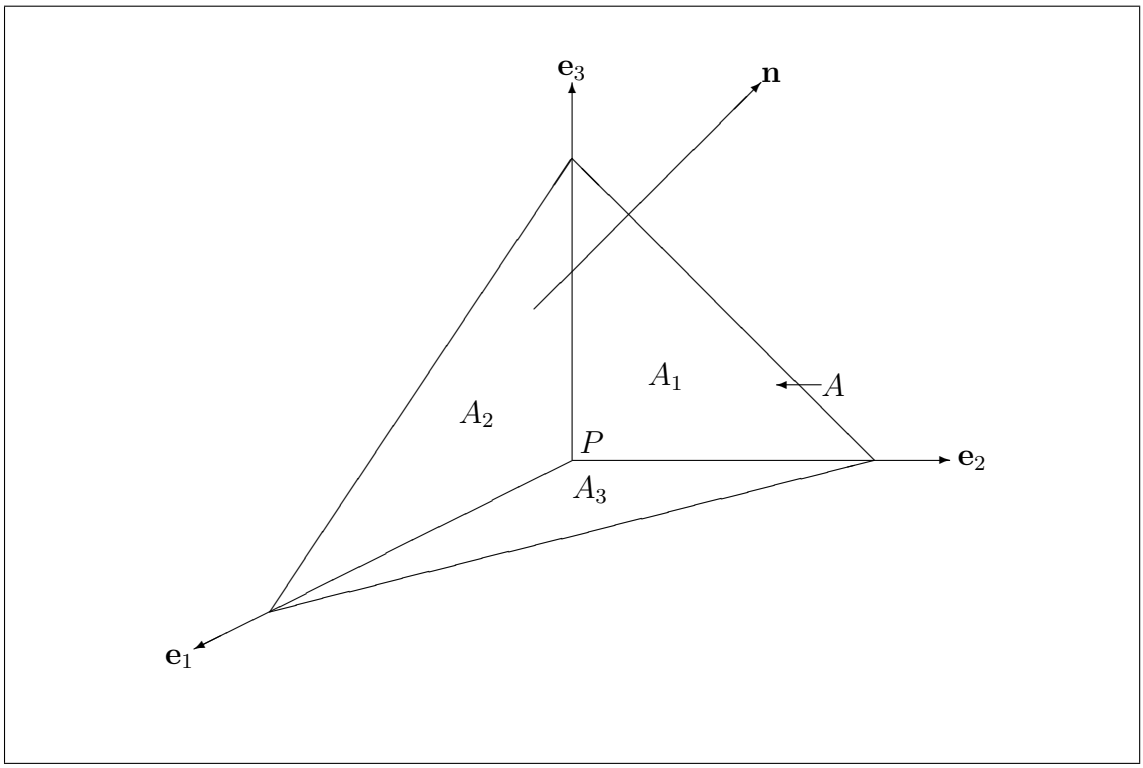

Figure 2.3: Tetrahedron formed by oblique plane $\mathbf{n}$ and coordinate planes $\mathbf{e_1}$, $\mathbf{e_2}$ and $\mathbf{e_3}$.

Since $V = V_1 + V_2$ and $S = S_1 + S_2$ we have

$$\mathbf{0} = \int_s [\mathbf{t}(\mathbf{x}, \mathbf{m}) + \mathbf{t}(\mathbf{x}, -\mathbf{m})] dS.$$

By an extension of this argument, a similar result holds for all sub-regions of s. We conclude

$$\mathbf{t}(\mathbf{x}, \mathbf{m}) = -\mathbf{t}(\mathbf{x}, -\mathbf{m}). \tag{2.4}$$

This shows that, at any given point, the stress vector acting on one side of a surface balances that on the other side.

Consider a tetrahedron (Figure 2.3), three faces of which meet perpendicularly at a point P and an oblique face whose unit normal $\mathbf{n}$ is at an arbitrary direction in the first quadrant and has an area of A. Let each perpendicular face be parallel to a Cartesian coordinate plane. Those planes with normals $\mathbf{e}_i$ have areas A_i. The angle between the face with normal $\mathbf{e}_i$ and the oblique face is $\cos^{-1}(n_i)$ and its area is related to A as $A_i = n_i A$. The volume of the tetrahedron is $\frac{1}{3}hA$, where h is the

distance of P from the oblique face.

We apply Eq.2.3 to the tetrahedron. Assuming that $|\rho d\mathbf{v}/dt|$ and $|\rho f(\mathbf{x})|$ are bounded with maximum values a and b, we take the limit $h \to 0$ keeping $\mathbf{n}$ fixed. Then, since for small h,

$$\left| \int_V \rho \frac{d\mathbf{v}}{dt} dV \right| \leq \frac{hAa}{3}$$

and

$$\left| \int_V \rho \mathbf{f}(\mathbf{x}) dV \right| \leq \frac{hAb}{3},$$

we have that, as $h \to 0$,

$$\frac{1}{A} \int_V \rho \frac{d\mathbf{v}}{dt} dV - \frac{1}{A} \int_V \rho \mathbf{f}(\mathbf{x}) dV \to \mathbf{0}. \tag{2.5}$$

Hence, dividing each term in Eq.2.3 by A and then replacing Eq.2.5 with the stress tensor

$$\lim_{h \to 0} \frac{1}{A} \int_S \mathbf{t}(\mathbf{x}, \mathbf{n}) dS \to \mathbf{0}.$$

The surface integral is taken over all four faces of the tetrahedron. In the limit the areas tend to zero, since the $A_i/A\,(i = 1; 2; 3)$ ratios are independent of h,

$$\mathbf{0} = \mathbf{t}(\mathbf{x}, \mathbf{n}) + \mathbf{t}(\mathbf{x}, -\mathbf{e_1}) \frac{A_1}{A} + \mathbf{t}(\mathbf{x}, -\mathbf{e_2}) \frac{A_2}{A} + \mathbf{t}(\mathbf{x}, -\mathbf{e_3}) \frac{A_3}{A}.$$

Applying Eq.2.4 and $A_i = n_i A$,

$$\mathbf{t}(\mathbf{x}, \mathbf{n}) = n_1 \mathbf{t}(\mathbf{x}, \mathbf{e_1}) + n_2 \mathbf{t}(\mathbf{x}, \mathbf{e_2}) + n_3 \mathbf{t}(\mathbf{x}, \mathbf{e_3}).$$

With

$$\begin{aligned} \mathbf{t}(\mathbf{x}, \mathbf{e_1}) &= \tau_{11}\mathbf{e_1} + \tau_{12}\mathbf{e_2} + \tau_{13}\mathbf{e_3} \\ \mathbf{t}(\mathbf{x}, \mathbf{e_2}) &= \tau_{21}\mathbf{e_1} + \tau_{22}\mathbf{e_2} + \tau_{23}\mathbf{e_3} \\ \mathbf{t}(\mathbf{x}, \mathbf{e_3}) &= \tau_{31}\mathbf{e_1} + \tau_{32}\mathbf{e_2} + \tau_{33}\mathbf{e_3} \end{aligned} \tag{2.6}$$

this becomes

$$\begin{aligned} \mathbf{t}(\mathbf{x}, \mathbf{n}) &= [n_1\tau_{11} + n_2\tau_{21} + n_3\tau_{31}]\mathbf{e_1} \\ &+ [n_1\tau_{12} + n_2\tau_{22} + n_3\tau_{32}]\mathbf{e_2} \\ &+ [n_1\tau_{13} + n_2\tau_{23} + n_3\tau_{33}]\mathbf{e_3} \\ &= t_1(\mathbf{x}, \mathbf{n})\mathbf{e_1} + t_2(\mathbf{x}, \mathbf{n})\mathbf{e_2} + t_3(\mathbf{x}, \mathbf{n})\mathbf{e_3}. \end{aligned} \tag{2.7}$$

In other words $t_i(\mathbf{x}, \mathbf{n}) = \tau_{ji} n_j$ where τ_{ij} are the components of the second order stress tensor.

Using the divergence theorem on the surface integral we obtain

$$\int_S \tau_{ji} n_j dS = \int_V \frac{\partial \tau_{ji}}{\partial x_j} dV.$$

This holds for all regions V assuming the integrand is continuous. In component form the total force (Eq.2.3) becomes

$$\frac{\partial \tau_{ji}}{\partial x_j} + \rho f_i = \rho \frac{dv_i}{dt}. \tag{2.8}$$

The density of angular momentum about the origin is defined to be $\mathbf{x} \times \rho\mathbf{v}$ per volume, so that the total angular momentum becomes

$$\int_V \mathbf{x} \times \rho\mathbf{v} dV. \tag{2.9}$$

Now the rate of change of angular momentum is equal to the total torque (Eq.2.9) around the origin

$$\frac{d}{dt} \int_V \mathbf{x} \times \rho\mathbf{v} dV = \int_S \mathbf{x} \times \mathbf{t}(\mathbf{x}, \mathbf{n}) dS + \int_V \mathbf{x} \times \rho\mathbf{f}(\mathbf{x}) dV, \tag{2.10}$$

where the surface integral is nonzero only when the stress vector $\mathbf{t}(\mathbf{x},\mathbf{n})$ has a nonzero component and direction perpendicular to both the surface normal vector $\mathbf{n}$ and the position $\mathbf{x}$ ($\mathbf{n}$ and $\mathbf{x}$ is not parallel). In component form

$$\frac{d}{dt} \int_V \epsilon_{ijk} x_j \rho v_k dV = \int_S \epsilon_{ijk} x_j t_k dS + \int_V \rho \epsilon_{ijk} x_j f_k dV,$$

where ϵ_{ijk} is the antisymmetric Levi-Civita symbol. Using the divergence theorem it follows that

$$\int_V \epsilon_{ijk} \left\{ \frac{\partial (x_j \tau_{rk})}{\partial x_r} + \rho x_j f_k - \rho \frac{d(x_j v_k)}{dt} \right\} dV = 0. \tag{2.11}$$

Now

$$\epsilon_{ijk} \frac{\partial (x_j \tau_{rk})}{\partial x_r} = \epsilon_{ijk} \left(\delta_{jr} \tau_{rk} + x_j \frac{\partial \tau_{rk}}{\partial x_r} \right) = \epsilon_{ijk} \left(\tau_{jk} + x_j \frac{\partial \tau_{rk}}{\partial x_r} \right)$$

and

$$\epsilon_{ijk} \frac{d(x_j v_k)}{dt} = \epsilon_{ijk} \left(v_j v_k + x_j \frac{dv_k}{dt} \right) = \epsilon_{ijk} \left(x_j \frac{dv_k}{dt} \right),$$

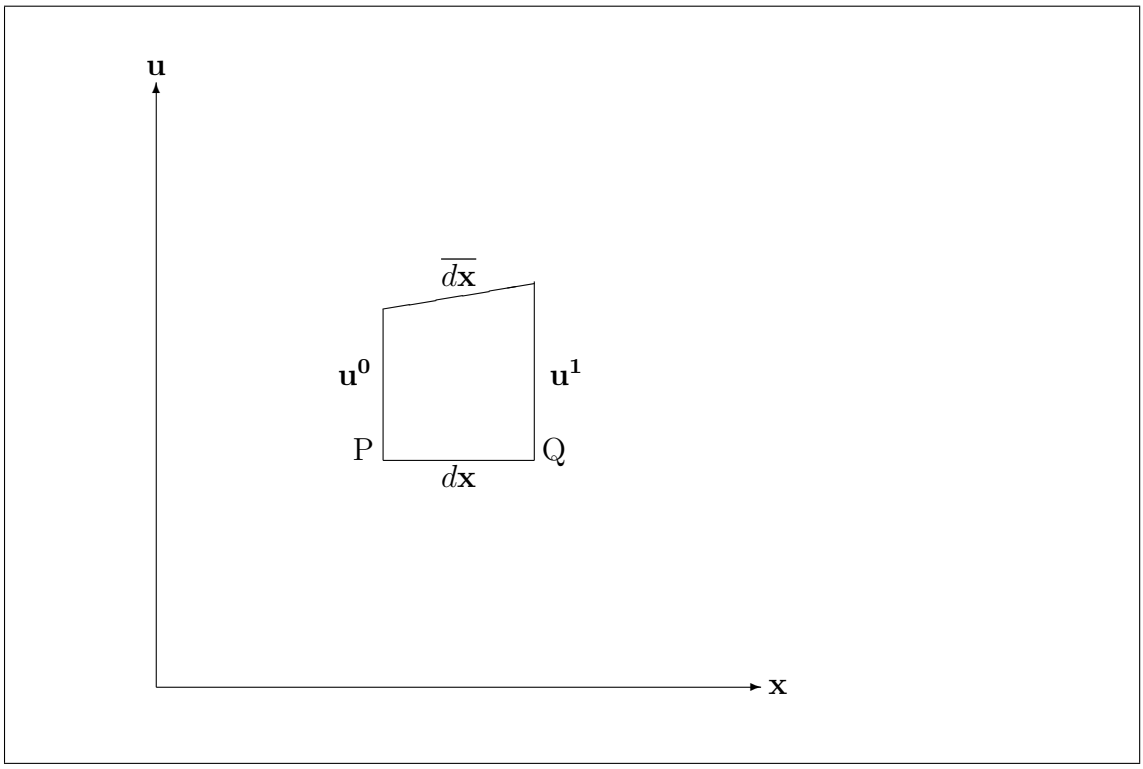

Figure 2.4: Deformations $\mathbf{u^0}$ for point P and $\mathbf{u^1}$ for Q at each end of a line element $d\mathbf{x}$.

then using the component-form total force

$$\int_V \epsilon_{ijk}\tau_{jk}dV = 0, \tag{2.12}$$

which hold for all regions V, then $\epsilon_{ijk}T_{jk} = 0$. Putting $i = 1, 2, 3$ in turn we find $\tau_{12} - \tau_{21} = 0$, $\tau_{13} - \tau_{31} = 0$, and $\tau_{23} - \tau_{32} = 0$. This gives a symmetric stress tensor

$$\tau_{ij} = \tau_{ji} \tag{2.13}$$

2.2 Strain

Using the notation of [4] consider (Figure 2.4) two material points P and Q at time $t = 0$, separated by a vector $d\mathbf{x}$, that move by displacements $\mathbf{u^0}$ and $\mathbf{u^1}$ respectively. We take the Taylor series of $\mathbf{u^1}$ around $\mathbf{u^0}$, with Einstein sum notation:

$$u_i^1 = u_i^0 + \frac{\partial u_i^0}{\partial x_j^0}(x_j^1 - x_j^0) + 0(x_j^1 - x_j^0)^2.$$

After dropping $0(x_j^1 - x_j^0)^2$

$$u_i^1 = u_i^0 + \frac{\partial u_i^0}{\partial x_j^0} dx_j. \tag{2.14}$$

The second term is a product of a second order tensor and a vector. This tensor can be separated into a symmetric and an anti-symmetric part,

$$u_i^1 = u_i^0 + \frac{1}{2}\left(\frac{\partial u_i^0}{\partial x_j^0} + \frac{\partial u_j^0}{\partial x_i^0}\right) dx_j + \frac{1}{2}\left(\frac{\partial u_i^0}{\partial x_j^0} - \frac{\partial u_j^0}{\partial x_i^0}\right) dx_j.$$

These last two terms refer to the strain tensor of second order

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i^0}{\partial x_j^0} + \frac{\partial u_j^0}{\partial x_i^0}\right) \tag{2.15}$$

and rigid-body rotation vector

$$\omega_j = \frac{1}{2}\left(\frac{\partial u_i^0}{\partial x_j^0} - \frac{\partial u_j^0}{\partial x_i^0}\right), \tag{2.16}$$

respectively. Thus

$$u_i^1 = u_i^0 + \varepsilon_{ij} dx_j + \epsilon_{ijk}\omega_j dx_k. \tag{2.17}$$

The second term on the right hand side is the rigid-body translation.

In the strain tensor the normal strains are ε_{ii}, $i = 1, 2, 3$ and the shear strains are $\varepsilon_{i\neq j}$, $i, j = 1, 2, 3$; while the symmetry of the tensor demands that $\varepsilon_{ij} = \varepsilon_{ji}$.

2.3 Hooke's Law

Introduced by the british physicist Robert Hooke (1635-1703), this is stated in component form as [4]

$$\tau_{ij} = c_{ijkl}\varepsilon_{kl}, \tag{2.18}$$

where c_{ijkl} are the components of the elastic tensor, and ε_{kl} is the strain tensor.

Due to the symmetry of both stress ($\tau_{ij} = \tau_{ji}$) and strain ($\varepsilon_{kl} = \varepsilon_{lk}$) tensors, the components of the fourth-order tensor have the following symmetries, respectively:

$$c_{ijkl} = c_{jikl} \quad , \quad c_{ijkl} = c_{ijlk}.$$

The stress is also related to what is called the strain-energy function $W = \frac{1}{2}c_{ijkl}\varepsilon_{ij}\varepsilon_{kl}$ which, combined with Hooke's law, is $\partial W/\partial\varepsilon_{ij} = \tau_{ij} = c_{ijkl}\varepsilon_{kl}$. This implies that

$$c_{ijkl} = c_{klij}$$

from $\partial^2 W/\partial\varepsilon_{ij}\partial\varepsilon_{kl} = \partial^2 W/\partial\varepsilon_{kl}\partial\varepsilon_{ij}$. These symmetries of the elastic tensor reduce the number of independent components from 81 to 21.

If an elastic property is the same in any direction, it is called isotropic [5]. This requires that the elastic tensor is not influenced by any rotation of a system of axes. In the following we are dealing with a isotropic system. With respect to a Cartesian reference system $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ the elastic tensor is c_{ijkl}, and with respect to the rotated system $(\mathbf{x'}_1, \mathbf{x'}_2, \mathbf{x'}_3)$ it is c'_{ijkl}. Then, because c_{ijkl} represents a 4th order tensor, the transformation becomes

$$c_{pqrs} = a_{ip}a_{jq}a_{kr}a_{ls}c'_{ijkl},$$

where $a_{mn} = \cos(\mathbf{x'}_m, \mathbf{x}_n) = \cos(\phi_{mn})$. Because of the invariance under rotation of the reference system, $c'_{ijkl} = c_{ijkl}$. Thus

$$c_{pqrs} = a_{ip}a_{jq}a_{kr}a_{ls}c_{ijkl}$$

which, as shown in [5], is satisfied only if

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu\delta_{ik}\delta_{jl} + \kappa\delta_{il}\delta_{jk}, \tag{2.19}$$

where we introduce the Lamé parameters λ, μ, and κ, as well as the Kronecker delta δ_{ij}.

Applying the symmetries of c_{ijkl} with respect to the two front and two back indices yields

$$[\kappa - \mu](\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) = 0.$$

Let $i = k$, $j = l$. Then, $\delta_{ik}\delta_{jl} = 9$, $\delta_{il}\delta_{jk} = 3$ is valid, and the last expression is only accurate if $\kappa - \mu = 0$ is true. Thus, the number of elastic constants for an isotropic medium has been reduced to two,

$$c_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{2.20}$$

Substituting this into Hooke's law and converting it using properties of the delta function we have the constitutive relationship between stress and strain:

$$\tau_{ij} = \lambda\delta_{ij}\varepsilon_{kk} + 2\mu\varepsilon_{ij}, \tag{2.21}$$

or in matrix form [3]:

$$\begin{pmatrix} \tau_{11} \\ \tau_{22} \\ \tau_{33} \\ \tau_{12} \\ \tau_{13} \\ \tau_{23} \end{pmatrix} = \lambda(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 2\mu \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{23} \end{pmatrix}.$$

Hooke's law for isotropic bodies is therefore given by Eq.2.21.

λ and μ are called Lamé parameters. They are constant for homogeneous materials and functions of space otherwise. These parameters are determined by experiments for any given material. Three of them are noteworthy here:

1. *Uni-axial tension:* A long, thin, cylindrical wire is stretched. When x_3 is taken along the wire, only the stress component τ_{33} is non-zero and related to strain ε_{33} through a scalar constant E to be derived below. Hooke's law predicts an extension along the direction of the tension and a contraction in the perpendicular direction. Thus

   $$\tau_{ii} = \lambda\varepsilon_{kk} + 2\mu\varepsilon_{ii} = 0,$$

   with $\tau_{ii} = 0$ for $i = 1, 2$ and $\tau_{33} = E\varepsilon_{33}$
   Subtracting the first two components from each other

   $$2\mu(\varepsilon_{11} - \varepsilon_{22}) = 0 \quad \rightarrow \quad \varepsilon_{11} = \varepsilon_{22}.$$

   On the other hand, by defining both ε_{11} and ε_{22} equal to $-\nu\varepsilon_{33}$, we have

   $$\lambda(1 - 2\nu)\varepsilon_{33} - 2\mu\nu\varepsilon_{33} = 0$$

$$\lambda = 2\lambda\nu + 2\mu\nu \quad \rightarrow \quad \nu = \frac{\lambda}{2(\lambda+\mu)}. \tag{2.22}$$

This is the ratio of the contraction to the elongation called Poisson's ratio.

From the equation for τ_{33} we have

$$\lambda(1-2\nu)\varepsilon_{33} + 2\mu\varepsilon_{33} = E\varepsilon_{33},$$

$$E = \lambda(1-2\nu) + 2\mu = \lambda\left[1 - \frac{\lambda}{\lambda+\mu}\right] + 2\mu,$$

$$E = \frac{(3\lambda+2\mu)\mu}{\lambda+\mu}. \tag{2.23}$$

This is the ratio of the tension and elongation known as Young's modulus

2. *Pure shear:* A bar with a rectangular cross section is in equilibrium under shearing forces in the x_1x_2-plane. $2\varepsilon_{12}$ may be interpreted as a decrease in the angle of two line elements that are parallel to the x_1 and x_2 axes respectively before the deformation. If $\varepsilon_{12} = \Gamma/2$, then the corresponding tension is given by $\tau_{12} = \mu\Gamma$.

   The ratio of τ_{12} to $2\varepsilon_{12}$ is known as the rigidity or shear modulus, i.e.

$$\mu = \gamma = \frac{\tau_{12}}{2\varepsilon_{12}}. \tag{2.24}$$

3. *Hydrostatic pressure:* A mass is placed in a large container with a liquid inside and a constant pressure p is applied to the liquid. This situation is also called a purely normal stress. By Pascal's law the body experiences only normal traction $\mathbf{T} = -p\mathbf{n}$. Thus, the stress is $\mathbf{S} = -p\mathbf{I}$ where $\mathbf{I}$ is the identity matrix. Here $\varepsilon_{11} = \varepsilon_{22} = \varepsilon_{33} \equiv \varepsilon$ and

$$\tau_{ij} = 3\lambda\varepsilon\delta_{ij} + 2\mu\varepsilon\delta_{ij}$$

$$\tau_{ii} = \tau = 3\lambda\varepsilon + 2\mu\varepsilon$$

$$p = (3\lambda+2\mu)\varepsilon.$$

   This gives $\varepsilon = p/(3\lambda+2\mu)$.

   The ratio of the tension to the expansion is known as three times the bulk modulus:

$$\frac{\tau}{\varepsilon} = 3\kappa,$$

where substituting for ε gives the bulk modulus:

$$\kappa = \lambda + \frac{2}{3}\mu. \tag{2.25}$$

We can now solve for the inverse of Hooke's law (i.e. for ε_{ij}):

$$\varepsilon_{ij} = \frac{1}{2\mu}\tau_{ij} - \frac{\lambda}{2\mu}\delta_{ij}\varepsilon_{kk}.$$

Through contraction ($i = j$, $\delta_{ii} = 3$) and renaming the index to k, we find that

$$\varepsilon_{kk} = \frac{\tau_{kk}}{3\lambda + 2\mu}.$$

Then replacing ε_{kk} back in the previous relation, we get

$$\varepsilon_{ij} = \frac{1}{2\mu}\tau_{ij} - \frac{\lambda}{2\mu(3\lambda + 2\mu)}\delta_{ij}\tau_{kk}. \tag{2.26}$$

In terms of Young's Modulus E and Poisson's Ratio ν, Hooke's Law and its inverse become

$$\tau_{ij} = \frac{E}{1+\nu}\left(\frac{\nu}{1-2\nu}\delta_{ij}\varepsilon_{kk} + \varepsilon_{ij}\right) \quad , \quad \varepsilon_{ij} = \frac{1}{E}\left[(1+\nu)\tau_{ij} - \nu\delta_{ij}\tau_{kk}\right]. \tag{2.27}$$

2.4 Primary and Secondary Waves

In the following treatment the notation of [1] and [2] is used.

The equation of motion for continuous media is

$$\rho\frac{\partial^2 u_i}{\partial t^2} = \tau_{ij,j} + f_{vi}, \tag{2.28}$$

where the comma notation indicates differentiation with respect to x_j, and repeated indices are summed according to Einstein's summation convention. It is satisfied at all points of a continuous medium.

As derived in [1], the constitutive relationship between stress and strain is:

$$\tau_{ij} = \lambda\Theta\delta_{ij} + 2\mu\varepsilon_{ij}, \tag{2.29}$$

where

$$\Theta = \nabla \cdot \mathbf{u} = u_{n,n}$$

and

$$\varepsilon_{ij} = [\nabla \mathbf{u} + (\nabla \mathbf{u})^T]/2 = (u_{i,j} + u_{j,i})/2.$$

Thus, more explicitly

$$\tau_{ij} = \lambda \delta_{ij} u_{n,n} + \mu(u_{i,j} + u_{j,i}),$$

with first derivatives

$$\tau_{ij,k} = \lambda \delta_{ij} u_{n,nk} + \frac{\partial \lambda}{\partial x_k} \delta_{ij} u_{n,n} + \mu(u_{i,jk} + u_{j,ik}) + \frac{\partial \mu}{\partial x_k}(u_{i,j} + u_{j,i}).$$

With no body forces and for non-homogeneous isotropic media, we have

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \tau_{ij,j} = \lambda \delta_{ij} u_{n,nj} + \frac{\partial \lambda}{\partial x_i} \delta_{ij} u_{n,n} + \mu(u_{i,jj} + u_{j,ij}) + \frac{\partial \mu}{\partial x_j}(u_{i,j} + u_{j,i}),$$

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \lambda u_{n,ni} + \mu(u_{i,jj} + u_{j,ij}) + \frac{\partial \lambda}{\partial x_i} u_{n,n} + \frac{\partial \mu}{\partial x_j}(u_{i,j} + u_{j,i}),$$

$$\begin{aligned} \rho \frac{\partial^2 u_i}{\partial t^2} &= \lambda(u_{1,1i} + u_{2,2i} + u_{3,3i}) + \frac{\partial \lambda}{\partial x_i}(u_{1,1} + u_{2,2} + u_{3,3}) \\ &+ \mu(u_{i,11} + u_{1,i1} + u_{i,22} + u_{2,i2} + u_{i,33} + u_{3,i3}) \\ &+ \frac{\partial \mu}{\partial x_1}(u_{i,1} + u_{1,i}) + \frac{\partial \mu}{\partial x_2}(u_{i,2} + u_{2,i}) + \frac{\partial \mu}{\partial x_3}(u_{i,3} + u_{3,i}), \end{aligned}$$

$$\rho \frac{\partial^2 u_i}{\partial t^2} = (\lambda + \mu) u_{n,ni} + \frac{\partial \lambda}{\partial x_i} u_{n,n} + \mu u_{i,nn} + \frac{\partial \mu}{\partial x_n}(u_{i,n} + u_{n,i}).$$

Then in vector short form in three dimensions:

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + \mu) \nabla(\nabla \cdot \mathbf{u}) + \nabla \lambda (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} + (\nabla \mathbf{u} + (\nabla \mathbf{u})^T) \nabla \mu. \tag{2.30}$$

Assuming λ and μ to be constant, and substituting the following vector calculus identity on the Laplacian

$$\nabla^2 \mathbf{u} = \nabla(\nabla \cdot \mathbf{u}) - \nabla \times (\nabla \times \mathbf{u})$$

into Eq.2.30 with λ and μ constant, gives

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + 2\mu) \nabla(\nabla \cdot \mathbf{u}) - \mu \nabla \times (\nabla \times \mathbf{u}). \tag{2.31}$$

Using the Helmholtz Ansatz the displacement $\mathbf{u}$ is decomposed as the sum of the gradient of a scalar potential and the curl of a vector potential:

$$\mathbf{u}(\mathbf{r}, t) = \nabla \Phi(\mathbf{r}, t) + \nabla \times \mathbf{A}(\mathbf{r}, t). \tag{2.32}$$

For any vector field $\mathbf{A}$, there is a zero divergence $\nabla(\nabla \times \mathbf{A}) = 0$
and for any scalar field Φ, there is a zero curl $\nabla \times (\nabla\Phi) = 0$.

Substituting the Helmholtz potentials (Eq.2.32) into the wave equation (Eq.2.31):

$$\rho\frac{\partial^2(\nabla\Phi + \nabla \times \mathbf{A})}{\partial t^2} = (\lambda + 2\mu)\nabla[\nabla \cdot (\nabla\Phi + \nabla \times \mathbf{A})] - \mu\nabla \times [\nabla \times (\nabla\Phi + \nabla \times \mathbf{A})]$$

$$\rho\frac{\partial^2(\nabla\Phi + \nabla \times \mathbf{A})}{\partial t^2} = (\lambda + 2\mu)\nabla[\nabla \cdot \nabla\Phi] - \mu\nabla \times [\nabla \times \nabla \times \mathbf{A}]$$

$$\rho\frac{\partial^2(\nabla\Phi + \nabla \times \mathbf{A})}{\partial t^2} = (\lambda + 2\mu)\nabla[\nabla^2\Phi] - \mu\nabla^2[\nabla \times \mathbf{A}],$$

where we used the given zero divergence and zero curl together with the following identities:

$$\nabla \times [\nabla \times \nabla \times \mathbf{A}] = -\nabla[\nabla \cdot (\nabla \times \mathbf{A})] + \nabla^2(\nabla \times \mathbf{A}) = \nabla^2(\nabla \times \mathbf{A})$$

and

$$\nabla \cdot (\nabla\Phi) = \nabla^2\Phi.$$

We group terms of the same potential, and pull out the gradient and curl operators:

$$\nabla\left[(\lambda + 2\mu)\nabla^2\Phi - \rho\frac{\partial^2\Phi}{\partial t^2}\right] = \nabla \times \left[\mu\nabla^2\mathbf{A} - \rho\frac{\partial^2\mathbf{A}}{\partial t^2}\right]. \tag{2.33}$$

These sides are equal for all t and $\mathbf{x}$, and also up to a constant which we take as zero. This allows us to separate the elastodynamic equation of motion into two partial differential equations:

$$\frac{\partial^2\Phi}{\partial t^2} = \frac{\lambda + 2\mu}{\rho}\nabla^2\Phi \tag{2.34}$$

for the longitudinal (P)rimary waves, and

$$\frac{\partial^2\mathbf{A}}{\partial t^2} = \frac{\mu}{\rho}\nabla^2\mathbf{A} \tag{2.35}$$

for the transverse (S)econdary waves, with

1. The P-wave Velocity $\quad c_P(r) = \sqrt{(\lambda + 2\mu)/\rho}$ $\quad$ and
2. The S-wave Velocity $\quad c_S(r) = \sqrt{\mu/\rho}$.

Primary or P-waves represent sound waves through solids and water. Secondary or S-waves can only move through solids since liquids and gases cannot support shear stresses. S-waves travel at typically half the speed of P-waves, but their amplitudes are usually much larger, leading to greater destruction during earthquakes.

Chapter 3

Research Procedures and Proposed Methods

In this chapter we use the following four methods. Firstly, the finite difference method, calculates derivatives of functions defined on a grid and helps to solve differential equations numerically. Secondly, the finite element method has been used only in the Gaussian feature physical setup. This approximates functions on domains where they are significantly different from zero. These domains are decomposed into square or rectangular elements. Each element is represented by fourth order base functions determined by coefficients on the corners and equidistant points in between. Thirdly, the Fourier Series is applied for only the single and multiple interface setups, to optimise the material property (MP) function to a model target MP function. The Fourier coefficients are used as the optimising parameters. Fourthly, an absorbing boundary condition is applied on numerical boundaries to get rid of the reflections there.

The MERLIN package contains optimisation methods that help to recover the model parameters to fit synthetic data or observed results.

3.1 Finite Difference Method

Derivatives of functions are numerically handled by the finite difference method [23, 24, 25], and can be used to solve the differential wave equation. This method is obtained by expanding any sufficiently smooth function $f(r)$ at $r = b$ to a Taylor series of 2nd order around $r = a$:

$$f(b) = f(a) + f'(a)h + \frac{f''(a)h^2}{2!} + O(h^3), \tag{3.1}$$

where primes indicate first and second derivatives, and $h = b - a$ should be within the range of convergence of this series. Functions are called smooth if they contain no discontinuities and there exist derivatives of up to infinite order.

Setting $a = 0$ and $b = h, 0, -h$:

$$\begin{aligned} f(h) &= f(0) + f'(0)h + f''(0)h^2/2, \\ f(0) &= f(0), \\ f(-h) &= f(0) - f'(0)h + f''(0)h^2/2, \end{aligned}$$

where $O(h^3)$ is dropped. In matrix notation

$$\begin{pmatrix} f(h) \\ f(0) \\ f(-h) \end{pmatrix} = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & -1 & \frac{1}{2} \end{bmatrix} \begin{pmatrix} f(0)h^0 \\ f'(0)h^1 \\ f''(0)h^2 \end{pmatrix}.$$

This can be solved for $f(0)h^0$, $f'(0)h^1$, and $f''(0)h^2$

$$f'(0) = \frac{f(h) - f(-h)}{2h} \tag{3.2}$$

$$f''(0) = \frac{f(h) - 2f(0) + f(-h)}{h^2}. \tag{3.3}$$

This is easily applied to space and time coordinates.

Using this in any of the partial differential equations given in the previous chapter with

$$\frac{\partial^2 u(\mathbf{r}, t)}{\partial t^2} = F(\mathbf{r}, t),$$

one can solve numerically the amplitude of the next time step given the amplitudes of the current and previous steps:

$$u_{n+1} = 2u_n - u_{n-1} + F(\mathbf{r}, t)\,\Delta t^2, \tag{3.4}$$

provided $F(\mathbf{r}, t)$, which is the spatial derivative part, is also expanded by this method.

3.2 Finite Element Method (FEM)

3.2.1 The Background

The method of finite elements was actually invented by structural engineers in the 1940s [26]. This method consists of subdividing the domain of the solution to a differential equation into elements of variable size. On each of the elements, the function is approximated by basis functions of a chosen order and their values (and possibly normal derivatives) must match up on the boundaries between elements to ensure continuity. Coefficients of these basis functions are evaluated from the solution on the vertices and typically on equidistant points within the element. These are then interpolated by the basis functions inside of the element. These elements are influenced only by their nearest neighbour elements that share a boundary (or corner) with them, and not by elements further away. However on the boundary of the domain Dirichlet, Neumann, or mixed boundary conditions are applied to make the solution of the partial differential equation unique. In our program we apply the FEM to the material parameters in order to minimise the objective function. Coefficients are calculated from this function even on the domain boundaries. Therefore the boundary condition is simple.

3.2.2 The Formalism

In what follows, we use the approach and notation from [27, 28]. Consider the problem of approximating a real-valued function $f(x)$ over a finite interval of the x-axis. Now we break up this interval into a number n of non-overlapping sub-intervals denoted $[x_i, x_{i+1}]$, $(i = 0, 1, 2, ..., n-1)$ and interpolate linearly between values of

$f(x)$ at the endpoints of each sub-interval. Then the piece-wise linear approximating function depends only on the values f_i of the functions at the nodal points x_i.

In a problem where $f(x)$ is given implicitly by a (differential, integral, functional, etc.) equation, the values f_i are the unknown parameters of the problem. In the problem of interpolation, the values f_i are known in advance.

In any one sub-interval, the appropriate linear approximating function is given by

$$p_1^{(i)}(x) = \alpha_i(x) f_i + \beta_{i+1}(x) f_{i+1}, \tag{3.5}$$

where

$$\alpha_i(x) = \frac{x_{i+1} - x}{x_{i+1} - x_i} \quad \text{and} \quad \beta_{i+1}(x) = \frac{x - x_i}{x_{i+1} - x_i}. \tag{3.6}$$

The local functions $\alpha_i(x)$ and $\beta_i(x)$ are known as *shape functions*. The values f_i and f_{i+1} are *nodal parameters*. There are only two parameters to an element, thus the element is said to have two *degrees of freedom*. The shape function is said to be of *first order*.

Hence the piecewise approximating function (Eq.3.5) over the whole interval $x_0 \le x \le x_n$ is

$$p_1(x) = \sum_{i=0}^{n} \varphi_i(x) f_i, \tag{3.7}$$

where

$$\begin{aligned} \varphi_0(x) &= \alpha_0(x) && x_0 \le x \le x_1 \\ &= 0 && x_1 \le x \le x_n \end{aligned}$$

$$\begin{aligned} \varphi_i(x) &= \beta_i(x) && x_{i-1} \le x \le x_i \\ &= \alpha_i(x) && x_i \le x \le x_{i+1} \\ &= 0 && x_0 \le x \le x_{i-1},\ x_{i+1} \le x \le x_n \end{aligned}$$

$$\begin{aligned} \varphi_n(x) &= \beta_n(x) && x_{n-1} \le x \le x_n \\ &= 0 && x_0 \le x \le x_{n-1} \end{aligned}$$

are pyramid functions, and represent an elementary type of *basis function*. These functions are identically zero except in the range $x_{i-1} \le x \le x_{i+1}$ and are

said to have *local support*.

Converting this into a standard coordinate $X = x/h - i$, $i = 0, 1, 2, ..., n$, with a uniform element size h:

$$\begin{aligned} \varphi_i(X) &= \beta_i(X) = 1 + X \quad -1 \leq X \leq 0 \\ &= \alpha_i(X) = 1 - X \quad 0 \leq X \leq 1 \\ &= 0 \qquad\qquad\quad X \leq -1,\ X \geq +1. \end{aligned}$$

This represents a standard linear basis function that is unity at point zero, and zero at plus or minus unity.

The function $f(x)$ can also be approximated to the second order by piecewise quadratic functions. They depend on f_i at the nodal points x_i, $i = 0, 1, 2, ..., n$, and intermediary points $x_{j+1/2}$, $i = 0, 1, 2, ..., n-1$. This time the piecewise approximating function (Eq.3.5) over the same interval is

$$p_2(x) = \sum_{i=0}^{n} \psi_i(x) f_i + \sum_{j=0}^{n-1} \chi_{j+1/2}(x) f_{j+1/2}, \tag{3.8}$$

where, in standard coordinates

$$\begin{aligned} \psi_i(X) &= 1 + 3X + 2X^2 \quad -1 \leq X \leq 0 \\ &= 1 - 3X + 2X^2 \quad 0 \leq X \leq 1 \\ &= 0 \qquad\qquad\quad X \leq -1,\ X \geq +1, \end{aligned}$$

which is unity at zero point, and zero at the other nodal and intermediary points,

$$\begin{aligned} \chi_{j+1/2}(X) &= 4X + 4X^2 \quad 0 \leq X \leq +1; \\ &= 0 \qquad\qquad X \leq 0,\ X \geq +1, \end{aligned}$$

which is unity at the intermediary point and zero at the nodal points, with $X = x/h - j$.

In this work we generalise only to fourth order approximations with

$$p_4(x) = \sum_{i=0}^{n} \psi_i(x) f_i + \sum_{j=0}^{n-1} [\chi_{j+1/4}(x) f_{j+1/4} + \chi_{j+1/2}(x) f_{j+1/2} + \chi_{j+3/4}(x) f_{j+3/4}], \tag{3.9}$$

where, in standard coordinates, $\psi_i(X)$ is always unity at zero and zero at any other point (nodal or intermediary), and $\chi_{j+1/4}(X)$, $\chi_{j+1/2}(X)$ and $\chi_{j+3/4}(X)$ is unity on only one intermediary point (1/4, 1/2 and 3/4, respectively) and zero at any other point. For the interpolation points in between both nodal or intermediary points, $\psi_i(X)$ with $\chi_{j+1/4}(X)$, $\chi_{j+1/2}(X)$ and $\chi_{j+3/4}(X)$ are otherwise non-zero and less than unity and represent the 4th order polynomial basis functions.

3.2.3 The Derivative Approximation

In general, first derivatives of piecewise approximating polynomials $p_1(x)$ and $p_2(x)$ are not the same as $f(x)$. We would now like to construct an approximating function which has the same values of function and first derivative as $f(x)$ at the nodal points x_i. These are piecewise cubic polynomials $p_3(x)$ such that

$$D^k f(x_i) = D^k p_3(x_i),\ (k = 0, 1;\ i = 0, 1, 2, ..., n),$$

where $D = d/dx$.

In the sub-interval $[x_i, x_{i+1}]$, the appropriate part of the cubic polynomials is given by

$$p_3^{(i)}(x) = \alpha_i(x) f_i + \beta_{i+1}(x) f_{i+1} + \gamma_i(x) f_i' + \delta_{i+1}(x) f_{i+1}', \tag{3.10}$$

where the shape functions are

$$\alpha_i(x) = \frac{(x_{i+1} - x)^2[(x_{i+1} - x_i) - 2(x - x_i)]}{(x_{i+1} - x_i)^3},$$

$$\beta_{i+1}(x) = \frac{(x - x_i)^2[(x_{i+1} - x_i) - 2(x_{i+1} - x)]}{(x_{i+1} - x_i)^3},$$

$$\gamma_i(x) = \frac{(x - x_i)(x_{i+1} - x)^2}{(x_{i+1} - x_i)^2},$$

$$\delta_{i+1}(x) = \frac{(x - x_i)^2(x - x_{i+1})}{(x_{i+1} - x_i)^2},$$

with $i = 0, 1, 2, ..., n - 1$ and f' denotes the derivative of f.

The piecewise approximating function (Eq.3.10) over the whole interval $x_0 \leq x \leq x_n$ is given by

$$p_3(x) = \sum_{i=0}^{n} [\varphi_i^{(0)}(x) f_i + \varphi_i^{(1)}(x) f_i'], \tag{3.11}$$

where $\varphi_i^{(0)}(x)$ and $\varphi_i^{(1)}(x)$ are easily obtained from the above $\alpha_i(x)$, $\beta_i(x)$, $\gamma_i(x)$ and $\delta_i(x)$.

3.2.4 The Bi-variate Approximation

We now consider approximation of a real-valued function $f(x, y)$ of *two* variables by piecewise continuous functions $p_1(x, y)$ over a bounded region R with boundary ∂R. The region is cut into a number of elements (which will only be of rectangular shapes in this thesis).

Let a typical rectangular element be $[x_i, x_{i+1}] \times [y_j, y_{j+1}]$, where $x_{i+1} - x_i = h_1 \, (0 \leq i \leq m-1)$ and $y_{j+1} - y_j = h_2 \, (0 \leq j \leq n-1)$. The bi-linear form interpolating $f(x, y)$ over the rectangular element is

$$p_1^{(i,j)}(x, y) = \alpha_{i,j}(x, y) f_{i,j} + \beta_{i+1,j}(x, y) f_{i+1,j} + \gamma_{i,j+1}(x, y) f_{i,j+1} + \delta_{i+1,j+1}(x, y) f_{i+1,j+1}, \tag{3.12}$$

where the shape functions are

$$\alpha_{i,j}(x, y) = \frac{(x_{i+1} - x)(y_{j+1} - y)}{h_1 h_2} \quad ; \quad \beta_{i+1,j}(x, y) = \frac{(x - x_i)(y_{j+1} - y)}{h_1 h_2};$$

$$\gamma_{i,j+1}(x, y) = \frac{(x_{i+1} - x)(y - y_j)}{h_1 h_2} \quad ; \quad \delta_{i+1,j+1}(x, y) = \frac{(x - x_i)(y - y_j)}{h_1 h_2}.$$

The piecewise approximating function (Eq.3.12) over $[x_0, x_m] \times [y_0, y_n]$ is given by

$$p_1(x, y) = \sum_{i=0}^{m} \sum_{j=0}^{n} \varphi_{i,j}(x, y) f_{i,j}. \tag{3.13}$$

The basis functions $\varphi_{i,j}(x, y)$ are identically zero except for the rectangular region $[x_{i-1}, x_{i+1}] \times [y_{j-1}, y_{j+1}]$.

For three-dimensional case, the method proceeds in a similar way.

3.2.5 The Implementation

The method is implemented for two and three dimensions here. The driver routine calculates the Lamé parameter as follows:

- Firstly, the following variables are defined:
  - The number of elements I for the x-axis, J for the y-axis and K for the z-axis of the domain in question, determining the number of rectangular elements.
  - The positions of the corners of adjacent boundaries $[x_{el}(i), y_{el}(j), z_{el}(k)]$ where $i = 0, 1, ..., I$; $j = 0, 1, ..., J$ and $k = 0, 1, ..., K$.
  - The input function $\lambda(x, y, z)$.
  - The final result, as variable splice, as output.
- Three index fields are created to facilitate the bookkeeping of FEM-coefficients, they are declared as two-dimensional arrays:

$$L_{il} = iN + l \quad \text{and} \quad M_{jm} = jN + m \quad \text{and} \quad N_{kn} = kN + n, \tag{3.14}$$

  where $l, m, n = 0, ..., N$.
- The coefficients $c_{IL,JM,KN}$ are calculated from the input function $\lambda(x, y, z)$ at the nodal points $(x_{el}(i), y_{el}(j), z_{el}(k))$, where $l = 0$, $m = 0$, and $n = 0$; and also at the equally spaced intermediary points, where $l \neq 0$, $m \neq 0$, and $n \neq 0$.
- Next it is determined in which element the grid point falls. Variables

$$t_x = \frac{x - x_{el}(i-1)}{x_{el}(i) - x_{el}(i-1)} \quad \text{and} \quad t_y = \frac{y - y_{el}(j-1)}{y_{el}(j) - y_{el}(j-1)} \quad \text{and} \quad t_z = \frac{z - z_{el}(k-1)}{z_{el}(k) - z_{el}(k-1)} \tag{3.15}$$

  local to that element are then calculated.
- The function is obtained as a linear combination of Nth order basis polynomials (here $N = 4$) in the element:

$$F(x, y, z) = \sum_{l=0}^{N} \sum_{m=0}^{N} \sum_{n=0}^{N} c_{iL,jM,kN}\, f_l^N(t_x) f_m^N(t_y) f_n^N(t_z), \tag{3.16}$$

where

$$f_n^N(t) = \prod_{n \neq n_0}^{N} \frac{Nt - n}{n_0 - n}. \tag{3.17}$$

3.3 Suppression of Reflections at the Boundary

One of the challenges in modelling wave physics on a numerical grid is to get rid of the back-reflection at the numerical boundaries. The wave ought to move off the mesh unhindered, or at least diminish there. One way this can happen is if its amplitude decrease with a numerical beach region at these boundaries [33]. These functions cannot completely wipe out the reflection, but the wave is simulated to attenuate as it approaches the boundary. In other words the propagation energy is absorbed by attenuating (damping) it. By analogy this can be interpreted by padding sponges to the walls of a studio to prevent sound waves from escaping to the outside.

In the literature [32] there is a space-based term of the form $\alpha(x)[\partial u/\partial x]$ with

$$\alpha(x) = a \exp(-c(x - w)^2)$$

on one boundary. We introduce the time-based term $\alpha(x)[\partial u/\partial t]$ to the differential equation as

$$\rho(x) \left[\frac{\partial^2 u}{\partial t^2} + 2\alpha(x) \frac{\partial u}{\partial t} \right] = \frac{\partial}{\partial t} \left[\lambda(x) \frac{\partial}{\partial t} u(x,t) \right]. \tag{3.18}$$

This *ad hoc* damping function is chosen as

$$\begin{aligned} \alpha(x) &= \alpha_0 \left[\frac{w - x}{w} \right] && \text{on the original boundary,} \\ &= \alpha_0 \left[\frac{x - d_o + w}{w} \right] && \text{on the opposite boundary,} \\ &= 0 && \text{in the middle,} \end{aligned}$$

and later, in squared form,

$$\begin{aligned} \alpha(x) &= \alpha_0 \left[\frac{w - x}{w} \right]^2 && \text{on the original boundary,} \\ &= \alpha_0 \left[\frac{x - d_o + w}{w} \right]^2 && \text{on the opposite boundary,} \\ &= 0 && \text{in the middle,} \end{aligned}$$

where d_o is the space mesh size, w is the zone inner boundary and α_0 defines the damping strength.

This method does not satisfactorily give the correct results of an absorption of the wave's reflected amplitude. That is, a reflection starts appearing at the inner boundary to the damping zone before the reflection at the numerical outer boundary could disappear. We have two reflections side-by-side, instead of none at all.

So far we have considered only methods that are all *ad hoc*. Other absorbing boundary conditions (ABC's) were tried. For these we go to [36] for a review. Here Givoli started off with a radiation condition given by Sommerfeld [40] for simple electromagnetic waves in one dimension. After defining the conditions for a non-reflecting boundary (the NRBC's), he went on to the two and three dimensions and introduced a technique of Engquist & Majda [41, 42] based on pseudodifferential equations. Then Lindman [43] suggested before anyone else at the time that one-way wave equations could be used as NRBC's. These were applied by Trefethen & Halpern [44, 45] using rational functions to approximate the irrational root function. Where [41, 42] used only Pade approximations for the rational function, [44, 45] considered Chebyshev, Chebishev-Pade, Newmann, L^2 and L^z approximations as well. Lindman [43] and Randall [46] considered developing NRBC's by first discretising the wave equations and then deriving boundary conditions with good transmitting properties with respect to the difference equations. In [41, 42] boundary conditions were developed in the continuous problem before discretisation of both the differential equations and their conditions. Wagatha [47] started off with the nonlocal condition of [41, 42] and then considered local approximations to this that depended on a parameter β. For certain choices of β, these NRBC's reduces to the local conditions of [41, 42]. Next we have Reynolds [48] who used a similar technique than [41, 42], with local NRBC's allowing the wave to continue over the boundary in cartesian coordinates and thus reducing the reflection coefficients there. This is a modified version of the one used in [42], although less rigorous. A scheme for Dirichlet/Newmann boundary conditions in elastodynamics was proposed by him. Bayliss

& Turkel [49, 50, 51] developed a sequence of boundary conditions, of increasing order, with axial and spherical symmetries and which is based on the asymptotic solutions at large distances. Feng [52] also obtained a sequence of NRBC's by first deriving an exact nonlocal integral relation on the boundary with an appropriate Green function. Then this is localised by an asymptotic approximation valid at large distances. Higdon [53, 54] obtained discrete NRBC's in a rectangular mesh by considering plane waves at discrete angles of perfect absorption to the mesh boundary. He then goes further and proves a theorem which states that if a NRBC is based on a rational symmetric approximation to the dispersion relation with respect to the outgoing waves (and cannot be improved by a simple modification of its coefficients), this would be equivalent to his BC's at suitable angles. This compares well with those of [41, 44, 47]. Earlier Keys [55] arrived at the same conditions and compared them nicely with [41, 48] and those of Lysmer & Kuhlemeyer [56] in the context of elastodynamics. Kriegsmann & Morawetz [57] obtained a NRBC for time-harmonic equations while Engquist & Halpern proposed a NRBC for the dispersion relation. And this concludes the review of NRBC's from [36]. Some of these methods, and newer ones, are explained in more recent books like [58, 59]. One new technique will now be explained next.

3.3.1 A perfectly matched layer.

Some of these NRBC's has been known to produce spurious reflections in certain conditions or couldn't satisfactorily reduce them. Then Berenger [38] proposed a new NRBC called the perfectly matched layer (PML) boundary condition in two-dimensional electromagnetic problems. This was later put in book form by him in [60].

In [38] two special cases are tried, the transverse-electric (TE) $[E_x, E_y, H_z]$ and the transverse-magnetic (TM) $[H_x, H_y, E_z]$ waves. In the TE problem we have three Maxwell equations and the magnetic field is split into the components of the electric field to increase these coupled equations to four. When the wave satisfy an

impedance condition $\sigma/\epsilon_0 = \sigma * /\mu_0$, where σ and $\sigma*$ are the electric and magnetic conductivities respectively, the medium is a vacuum and the wave propagates unhindered through the barrier without a reflection. Other special cases are considered. If $\sigma_x = \sigma_y = \sigma*_x = \sigma*_y = 0$, where subscripts x and y indicates the components of conductivities, then we have Maxwell equations for a vacuum. If $\sigma_x = \sigma_y \neq 0$ and $\sigma*_x = \sigma*_y = 0$, the medium is conducting the waves. If $\sigma_x = \sigma_y \neq 0$ and $\sigma*_x = \sigma*_y \neq 0$, the medium is absorbing the waves. When $\sigma_y = \sigma*_y = 0$, a wave propagating along the x-axis, (E_y, H_{zx}), is absorbed, but not waves (E_x, H_{zy}) traveling in the y-direction. Vice versa when $\sigma_x = \sigma*_x = 0$.

When we approach the TM problem, however, it is the electric field that must split into the components of the magnetic field. And similar conditions apply here than in the TE problem. After all the derivations and calculations of Maxwell's equations, we finally arrive at a solution wave form of $\psi(\rho) = \psi(0)\sqrt{R(\rho.\theta)}$ where $R(\rho, \theta) = e^{-2f(\theta)\rho}$ which impinges the absorbing layer at angle θ.

3.3.2 A New Non-Reflecting Boundary Method

A better method is desirable and has been presented and published in [37]. It is a theory of non-reflecting boundary conditions (NRBC) where we use an exponential function of an absorbing 3rd order polynomial.

In [37] the above-mentioned exponential function is $e^{\int_0^x \alpha(x)dx}$ where $\alpha(x)$ is the quadratic polynomial above. Here we define $F(x) = \int_0^x \alpha(x)dx$ for simplification. Any constants resulting from the integration are absorbed in the constant F_o when $F(x)$ is restated below. Then $\alpha(x) = dF(x)/dx$ and $d\alpha(x)/dx = d^2F(x)/dx^2$. $F(x)$ has the following properties:

$$F(x) = 0, \qquad x < w$$

$$\lim_{x\to\infty} F(x) = \infty.$$

We then have $v(x-t) = e^{F(x)}u(x-t)$ with time and space first derivatives

$$\dot{u} = e^{-F}\dot{v} = -e^{-F}v' \qquad \text{and} \qquad u' = (e^{-F}v)' = -e^{-F}F'v + e^{-F}v' = -F'u - \dot{u},$$

where $u' = \partial u/\partial x$, and $\dot{u} = \partial u/\partial t$.

Without using $\lambda(x)$ or $\rho(x)$ we have $\ddot{v}(x,t) = v''(x,t)$. When we substitute for $v(x,t)$ in this form, the differential equation becomes

$$\ddot{u} + 2F'\dot{u} = u'' - [(F')^2 - F'']u \qquad \text{or} \qquad \ddot{u} + 2\alpha\dot{u} = u'' - [\alpha^2 - \alpha']u, \tag{3.19}$$

where now

$$\begin{aligned} F(x) &= F_0 \left[\frac{x-w}{d-w}\right]^3 && x \in [w,d] \\ &= 0 && \text{otherwise,} \end{aligned}$$

with f_0 arbitrary, and

$$\begin{aligned} \alpha(x) &= \alpha_o \left[\frac{x-w}{d-w}\right]^2 && x \in [w,d] \\ &= 0 && \text{otherwise,} \end{aligned}$$

where $\alpha_o = 3F_0/(d-w)$.

With uniform λ implemented, w is set to $d/2$. With $\lambda(x)$ as defined in section 4.2 by its function and b_r at $d/2$, there are 2 regions, one on each boundary. Then

$$\begin{aligned} F(x) &= F_0 \left[\frac{x-w_2}{d-w_2}\right]^3 && x \in [w_2,d] \\ &= F_0 \left[\frac{w_1-x}{w_1}\right]^3 && x \in [0,w_1] \\ &= 0 && \text{otherwise,} \end{aligned}$$

where w_1 and w_2 are close to the initial and opposite numerical boundary respectively.

Here again, for $v(x-ct) = e^{F(x)}u(x-ct)$, the first derivatives are

$$\dot{u} = e^{-F}\dot{v} = -e^{-F}v'c \qquad \text{and} \qquad u' = (e^{-F}v)' = -e^{-F}F'v + e^{-F}v' = -F'u - \dot{u}/c.$$

Using $\lambda(x)$ and $\rho(x)$ we get $\rho(x)\ddot{v}(x,t) = [\lambda(x)v'(x,t)]'$. The following differential equation is obtained:

$$\frac{\ddot{u}}{c^2} + \left[\frac{b^2}{c^2} + 2F'\right]\frac{\dot{u}}{c} = u'' - \{(F')^2 - F''\}u \tag{3.20}$$

or

$$\frac{\ddot{u}}{c^2} + \left[\frac{b^2}{c^2} + 2\alpha\right]\frac{\dot{u}}{c} = u'' - \{\alpha^2 - \alpha'\}u, \tag{3.21}$$

with $\alpha(x)$ and where $c^2(x) = \lambda(x)/\rho(x)$ and $b^2(x) = \lambda'(x)/\rho(x)$.

When $\lambda(x)$ is uniform and set to unity ($\lambda'(x)$ is zero), then $c(x) = 1$ and $b(x) = 0$, reducing this differential equation to the previous differential equation derivation.

This is producing satisfactory results where the wave $u(x,t) = e^{-F(x)}v(x,t)$ decays almost completely on entering the damping zone with no reflections elsewhere in the zone.

In [37] similar derivations were done and we implemented them in two and three dimensions with similar results, where the first and second space derivatives were replaced by their gradient and Laplacian equivalents respectively.

3.4 Fourier Series

For optimisation purposes we approximate the $\lambda(\mathbf{x})$ by a truncated Fourier series. The coefficients of this Fourier series are then used as Merlin minimisation parameters.

The Fourier series decomposes a given periodic function or signal into a sum of simple oscillating functions, namely sines and cosines (or complex exponentials).

For a 2π-periodic, continuous function $f(x)$ that is integrable on $[\pi, \pi]$, the numbers

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)\,dt, \quad n \geq 0$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(nt)\,dt, \quad n \geq 1$$

are called the Fourier coefficients of f. One introduces the partial sums of the Fourier series for f, often denoted by

$$(S_N f)(x) = \frac{a_0}{2} + \sum_{n=1}^{N} [a_n \cos(nx) + b_n \sin(nx)], \quad N \geq 0.$$

The partial sums for f are called trigonometric polynomials. One expects that the functions $S_N f$ approximate the function f, and that the approximation improves as N tends to infinity on points of continuity only. The infinite sum

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)] \tag{3.22}$$

is called the Fourier series of f.

However in our application we do not use the cosine terms because they do not agree with the required boundary conditions.

3.5 The Merlin Minimisation Package

For minimisation we use the freely available package MERLIN [15, 16]. This package contains a colection of optimisation techniques to run on our program and requires two dimensions (number of terms M, number of parameters N), the input parameters, and one of two user supplied subroutines: `Subsum` for the Sum-of-Square minimisation of the objective function, or `Funmin` for the minimisation of a general function.

Suppose we have the points (t_i, d_i); $i = 1, 2, ..., M$, where M is the 'number of terms' of the objective/general function. We want to construct a function $G(t)$ such that $d_i \approx G(t_i)$ for all $i = 1, 2, ..., M$. We model $G(t) = G(t, m_1, m_2, ..., m_N)$ with input parameters m_j; $j = 1, 2, ..., N$, where N is the 'number of parameters' which will be varied to achieve the above goal. Equivalently this corresponds to minimising the objective function

$$F(m_1, m_2, ..., m_N) = \sum_{i=1}^{M} f_i^2(m_1, m_2, ..., m_N) = \sum_{i=1}^{M} [d_i - G(t_i, m_1, m_2, ..., m_N)]^2$$

where d_i are values from the target function and $\mathbf{m} = (m_1, m_2, ..., m_N)$ is the parameters, in vector form, each component optionally subject to lower and upper limits $[l_i, u_i]$. Each and every parameter can also be constrained (or fixed) individually to certain values and thus not take part in the minimisation process.

When an optimised technique is chosen to run the program, MERLIN makes use of the gradient, Jacobian, and Hessian of $F(\mathbf{m})$ to calculate the next step in bringing the point $\mathbf{m}$ closer to the minimum position $\mathbf{m}_0$ in the parameter space. These are

$$g = \nabla F, \qquad J_{ij} = \frac{\partial F_i}{\partial a_j}, \qquad H_{ij} = \frac{\partial^2 F}{\partial a_i \partial a_j}.$$

In this thesis the Levenberg-Marquardt algorithm is used and only the gradient and Jacobian is relevant (see section 3.5.3 below), thus the process terminates when these quantities reach a minimum tolerance value, whichever comes first. Our above goal is then reached when $\mathbf{m} \approx \mathbf{m}_0$ and $F(\mathbf{m}_0) \approx 0$. The technique also terminates when the rate of change in the parameters or in the objective function passes below a minimum tolerance each, separately. Then it also terminates after a maximum number of steps (calls or iterations) have been reached and it returns a message saying that no further progress is possible. But in this thesis all tolerances is set to zero.

In these optimisation techniques the input parameters $\mathbf{m}$ are initialised to start the process. No intervals are set on, and no equality or inequality constraints apply to the parameters (or the objective function) in this thesis. Thus no regularation is done here. These parameters are built into the material property function through $G(t, \mathbf{m})$ in question. Correspondingly, an analytic function of the material properties is built for $\mathbf{d}$, to which the optimised function must fit after the technique is run. A procedure is used to derive the seismic signals (wave equations, amplitude) via finite differences in both cases: one from the analytically given target, and the other from the optimised function (where the parameters are yielded by MERLIN). The target function must return a reference or calculated signal $\mathbf{d}$ and the optimised function is initialised to represent the modeled signal $G(t, \mathbf{m})$. As shown below the objective function is defined as a sum over propagation time of the squared difference between the modeled (measured) and referenced (calculated) signals. When the technique runs, these parameters must return a set of values (the minimum point $\mathbf{m}_0$) such that the optimised function must closely fit the model and the objective function approaches zero.

3.5.1 User supplied subroutines

In this thesis we use `Subsum`. The objective function for one dimension is then a sum of squared terms over the time grid $t_k, k = 1, ..., K$, where K is the number of terms, and at a fixed space measurement point x_m:

$$F(t) = \sum_{k=1}^{K} f_k(t_k)^2 = \sum_{k=1}^{K} [w(x_m, t_k) - u(x_m, t_k)]^2, \tag{3.23}$$

where $G(t_k, \mathbf{m}) = w(x_m, t_k)$ is the amplitude resulting from the model (optimised function) and $d_k = u(x_m, t_k)$ the amplitude resulting from the simulated target function.

In two dimensions and three dimensions we use several points (x_m, y_m) or (x_m, y_m, z_m), $m = 1, ..., M$, to measure the amplitude. The origin (a_x, a_y) or (a_x, a_y, a_z) of the point-source, outward-moving wave is one of those points. These are described in chapter 4. The objective function is now summed not only over the time grid but also over these measurement points:

$$F(\mathbf{r}, t) = \sum_{k=1}^{K} \sum_{m=1}^{M} f_{km}(\mathbf{r}_m, t_k)^2 = \sum_{k=1}^{K} \sum_{m=1}^{M} [w(\mathbf{r}_m, t_k) - u(\mathbf{r}_m, t_k)]^2, \tag{3.24}$$

where now m is the current point index from a total of M measurement points.

First, a reference amplitude $u(\mathbf{r}_m, t)$ is calculated during `Subsum`'s first run to obtain synthetic data. This is done using the material parameters defined by the target functions. Then $w(\mathbf{r}_m, t)$ is calculated using the material parameters as determined by the model input parameters through MERLIN. The objective function should then decrease to zero, when $w(\mathbf{r}_m, t)$ approach $u(\mathbf{r}_m, t)$, as the calculated material parameters approach the original target material property.

3.5.2 The Minimisation Algorithms

The minimisation techniques used in Merlin are the following: The quasi-newton methods dfp and bfgs described in [18] and [19] respectively, the conjugate gradient method congra described in [20], the geometric method simplex in [21] and the Levenberg-Marquard method leve in [22]. See also [11, 13, 14] for further explana-

tions of these methods.

The first three algorithms are gradient methods that involve first (gradient) and second (Hessian) derivatives of the objective function (of quadratic form) with respect to the input parameters to calculate a search direction and update the parameters accordingly. The steepest-descent method updates the parameters using the gradient of the objective function. In the conjugate-gradient method the search direction must first be adjusted from the gradient to satisfy conjugate conditions. Among the many different methods to alter the search direction are Fletcher-Reeves (FR), Polak-Ribière (PR) and Hestenes-Stiefel (HS). These methods are equivalent for quadratic function problems but differ from each other otherwise. The minimisation technique in Merlin with these options is congra.

A more straightforward method of calculating these quadratic functions is the Newton search method. Here the space metric information is calculated from the Hessian of the objective function whose inverse is multiplied by the gradient, and then used to update the parameters. An approximation to this is what is called the quasi-Newton methods, in which all second and higher order derivatives within the Hessian are dropped.

Sometimes it is necessary to precondition the gradient before it is used in the update. In variable-metric methods the preconditioning operator is allowed to approach the inverse Hessian by varying it from one iteration to the next. It starts by behaving like the steepest-descent or conjugate-gradient methods and ends through the behaviour of the Newton or quasi-Newton methods. The complementary dfp and bfgs are two such methods used in Merlin. This also suggests a Broyden family of methods in which a mixture of these methods is used together, one weighted with the other [14].

Simplex is a technique based on imposing a two-dimensional equilateral triangle or three-dimensional equidistant tetrahedron on the geometry of the problem. Then

this simplex moves and changes shape by reflection and contraction until a minimum of the function is reached.

3.5.3 Levenberg-Marquardt method

Currently we use the leve minimisation subroutine in Merlin: The Levenberg-Marquardt (LM) algorithm leve is a restricted-step method, only in the L_2-norm for least-square non-linear problems, that locates a minimum of a function expressed as the sum of squares of nonlinear functions. According to the abstract of Lourakis [22]: "[It] can be thought of as a combination of [the] steepest-descent and the Gauss-Newton method[s]." When the current solution is far from the correct one, the steepest-descent behaviour dominates: slow but guaranteed to converge. When close to the correct solution, the Gauss-Newton behaviour takes over.

In the following, vectors are small boldface and A^T is the transpose of matrix A. $||\cdot||$ is a 2-norm. We map a *parameter vector* $\mathbf{p} \in \mathbb{R}^m$ to an estimated *measurement vector* $\hat{\mathbf{x}} \in \mathbb{R}^n$ with an assumed function $\hat{\mathbf{x}} = f(\mathbf{p})$. An initial parameter estimate $\mathbf{p_0}$ and corresponding measurement $\mathbf{x}$ is provided. It is desired to find a vector $\mathbf{p}_+$ that best satisfies the functional relation f, i.e., that minimises the squared distance $\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon} = \mathbf{x} - \hat{\mathbf{x}} = \boldsymbol{\delta}_x$. The basis of LM is a linear approximation to f in the neighbourhood of $\mathbf{p}$. For a small $||\boldsymbol{\delta}_p||$ a Taylor series expansion leads to the approximation:

$$f(\mathbf{p} + \boldsymbol{\delta}_p) \approx f(\mathbf{p}) + J\boldsymbol{\delta}_p, \tag{3.25}$$

where J is the Jacobian matrix of $f(\mathbf{p})$. Like all non-linear optimisation methods, LM is iterative, starting from $\mathbf{p_0}$ it produces a series of vectors $\mathbf{p_1}, \mathbf{p_2}, ...$ that converge to the local minimiser $\mathbf{p}_+$ for f. At each step it is required to find the $\boldsymbol{\delta}_p$ that minimises the following:

$$||\mathbf{x} - f(\mathbf{p} + \boldsymbol{\delta}_p)|| \approx ||\mathbf{x} - f(\mathbf{p}) - J\boldsymbol{\delta}_p|| = ||\boldsymbol{\epsilon} - J\boldsymbol{\delta}_p||.$$

The desired $\boldsymbol{\delta}_p$ is therefore a solution to a linear least-square problem: the minimum is obtained when $J\boldsymbol{\delta}_p - \boldsymbol{\epsilon}$ is orthogonal to the column space of J. This leads to $J^TJ(\boldsymbol{\delta}_p - \boldsymbol{\epsilon}) = \mathbf{0}$. This yields $\boldsymbol{\delta}_p$ as a solution to the so-called *Normal*

equation:

$$J^T J\boldsymbol{\delta}_p = J^T\boldsymbol{\epsilon}. \tag{3.26}$$

The matrix $J^T J$ on the left hand side of Eq.3.26 is an approximate Hessian. The LM actually solves a variation of Eq.3.26, known as the *augmented normal equations*:

$$N\boldsymbol{\delta}_p = J^T\boldsymbol{\epsilon}, \tag{3.27}$$

where $N = J^T J + \nu \text{ diag}(J^T J)$. The strategy of adjusting diagonal elements of N is called *damping* and the ν is referred to as the *damping term*.

If the updated parameter term leads to a reduction in the error $\boldsymbol{\epsilon}$, the update is accepted and the process repeats with a decreased value of ν. Otherwise ν is increased, the augmented normal equation is solved again, and the process iterates until a value of $\boldsymbol{\delta}_p$ is found that decreases error. The process of repeatedly solving Eq.3.27 for different values of the damping term until an acceptable parameter vector update is found, corresponds to one iteration of the LM algorithm.

If the damping term is set to a larger value, the matrix N is nearly diagonal and the LM update step $\boldsymbol{\delta}_p$ is near the steepest-descent direction. The magnitude of $\boldsymbol{\delta}_p$ is reduced contributing to its slowness in this behaviour. Damping also handles situations where the Jacobian is rank-deficient and $J^T J$ is singular. The LM then defensively navigates a region of the parameter space where the model is highly non-linear. If damping is small, the LM step approximates the exact quadratic step appropriate for a fully linear problem in a Gauss-Newton way. LM is adaptive because it controls its own damping: it raises damping if a step fails to reduce error, otherwise damping is reduced. In this way the LM is capable of alternating between a slow descent approach when far from the minimum and a fast convergence when in the neighbourhood of the minimum.

The LM algorithm terminates when at least one of the following conditions is met:

- The magnitude of the gradient of $\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}$, i.e. $||J^T\boldsymbol{\epsilon}||$ drops below a threshold ε_1
- The relative change in the magnitude of $\boldsymbol{\delta}_p$ drops below a threshold ε_2

- The error $\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}$ drops below a threshold ε_3
- The maximum number of iterations k_{max} is completed.

Chapter 4

Results

We used a laptop with good computing (1GHz CPU) and memory (1GB RAM) capability for the simulations. In one dimensional calculations the program run-times are between seconds and not more than about 3 minutes. In two dimensions the calculations take between a few minutes and about an hour to complete. In three-dimensional case, however, the program runs for hours, even up to half a day. For three dimensions or more complicated setups, these calculations then become very time-consuming to complete. A supercomputer with more memory capacity or a server with more than one processor for parallel computing is needed to achieve faster convergence and shorter run-time.

Here we discuss damping conditions at the boundaries and the behaviour of the optimisation method before giving material property results for the following setups. In one dimensional setup are the details firstly for one interface and secondly for more than one interfaces. In two dimensional setup we have similar results for one interface and then parallel multiple interfaces. All the results for the interfaces made use of the Fourier series method. However, all results for the Gaussian feature used the finite element method (FEM) instead. So for two and three dimensions we have comparative results for a Gaussian feature standing alone without any interfaces present. In three dimensions there are no data for any interface setups. These results are more clearly shown in the material property plots. With the property plots, we also give plots of the wave propagation.

4.1 Preliminaries

4.1.1 Damping Condition

Throughout the calculations in sections 4.2, 4.3 and 4.4 the following damping conditions are in effect.

One dimension

There is only one zone on the original side of the mesh, with the zone inner boundary a tenth of the way from the mesh boundary: $w = 0.1d$, where the total distance is d and fixed to 1. The zone functions $f(x)$ are cubic polynomials as defined in section 3.3.2. The strength of each zone increases from zero at the inner boundary up to the value of $f_0 = 5$ on the outer side.

The damping zone function (Figure 4.2) is restated here:

$$\begin{aligned} f(x) &= f_0 \left[\frac{w - x}{w} \right]^3 && x \in [0, w] \\ &= 0 && \text{otherwise.} \end{aligned}$$

This is done so that the reflected waves, sampled at position $x = 0.125$, can also assist in optimising the objective function at later propagation times $t \to 2$. This is seen in figures 4.6 and 4.14 in section 4.2. Since the optimisation does a good job with these reflections, we keep these one dimensional results. They are however absorbed at the original boundary where damping is still in place.

Two and three dimensions

In one dimension the boundary conditions apply to the amplitude $\mathbf{u}(x, t)$ used as the function in the differential wave equation below. However, in two and three dimensions the calculated function for the partial differential equation is the scalar potential $\mathbf{\Phi}(x, t)$, and the same boundary conditions apply to that.

The damping function on all sides in two and three dimensions is

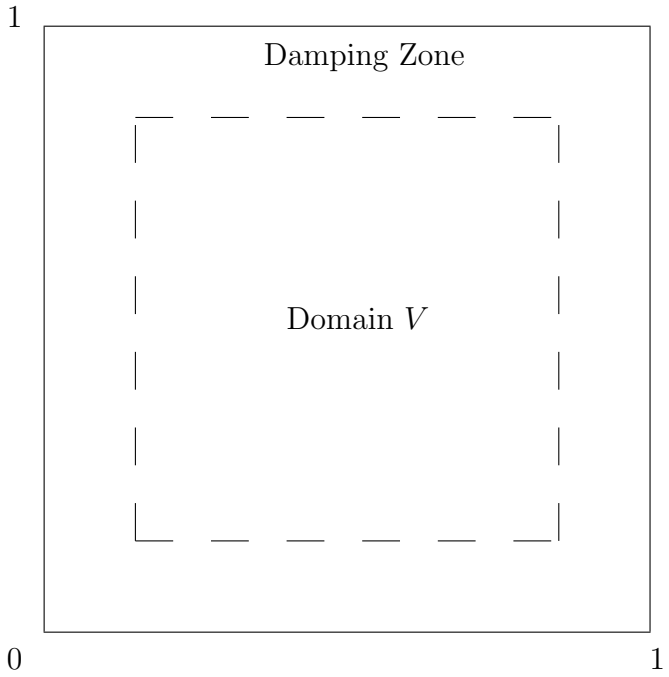

Figure 4.1: The domain V with the damping zones at the sides in 2D. In 3D this becomes a cube of work volume within a cube of domain V.

$$
\begin{aligned}
f_1(x) &= f_0 \left[\frac{w_{x_1} - x}{w_{x_1}} \right]^3 && x \in [0, w_{x_1}] \\
&= f_0 \left[\frac{x - w_{x_2}}{d - w_{x_2}} \right]^3 && x \in [w_{x_2}, d] \\
&= 0 && \text{otherwise,} \\
f_2(y) &= f_0 \left[\frac{w_{y_1} - y}{w_{y_1}} \right]^3 && y \in [0, w_{y_1}] \\
&= f_0 \left[\frac{y - w_{y_2}}{d - w_{y_2}} \right]^3 && y \in [w_{y_2}, d] \\
&= 0 && \text{otherwise,} \\
f_3(z) &= f_0 \left[\frac{w_{z_1} - z}{w_{z_1}} \right]^3 && z \in [0, w_{z_1}] \\
&= f_0 \left[\frac{z - w_{z_2}}{d - w_{z_2}} \right]^3 && z \in [w_{z_2}, d] \\
&= 0 && \text{otherwise.}
\end{aligned}
\quad .
$$

Then

$$f(\mathbf{r}) = f_1(x) + f_2(y) + f_3(z),$$

where the inner positions are $w_{x_1} = w_{y_1} = w_{z_1} = 0.15$ and $w_{x_2} = w_{y_2} = w_{z_2} = 0.85$, and the zone strength $f_0 = 10$. Zones x_1 and x_2 add constructively with both

y_1 and y_2, and z_1 and z_2 at the corners and sides. When working in two dimensions, only $f_1(z)$ and $f_2(z)$ are used. See Fig.4.3.
These zone strength and position values (here and in one dimension) are chosen large enough until reflection of the waves can no longer be seen.

As can be seen in the plots of wave propagation in two and three dimensions, the amplitude falls to zero a short distance from and parallel to the boundaries. Those regions of near-zero amplitude marks the damping area. Compared with the amplitude in the middle region, one can see the effectiveness of damping at the corners, sides and faces. Thus, no boundary reflections are seen in these plots.

4.1.2 Objective Function

Figure 4.4 (for one-dimensional multi-layered results below) shows how the sum-of-squares was minimised by the Levenberg-Marquard (LM) technique with a maximum iteration setting after which the MERLIN package would stop with the message indicating that fact. In most cases it stops below its maximum setting, where the message would be that no further progress is possible.

For the optimisation process to proceed a gradient $\nabla_{\mathbf{p}} f(\mathbf{p})$ of the objective function $f(\mathbf{p})$ with respect to the input parameter space $\mathbf{p}$ is calculated. There is one gradient component for each dimension of the parameter space. Say one has N input parameters, p_n, $n = 1, .., N$, that represents the N-dimensional parameter space. Each component $\partial f(\mathbf{p})/\partial p_n$ of this gradient is calculated one MERLIN call at a time. The technique would calculate one component of the gradient and advance to the next call, do the next component and advance, until all the components have been done in about N steps before $f(\mathbf{p})$ can move to its next (lower) value. As a result a plateau is formed in the objective function data around every N, each one successively lower than the previous one but constant during the evaluation of the whole gradient. Once the gradient is evaluated, it is used to move the objective function to a new (and lower) value. The gradient calculation starts all over again

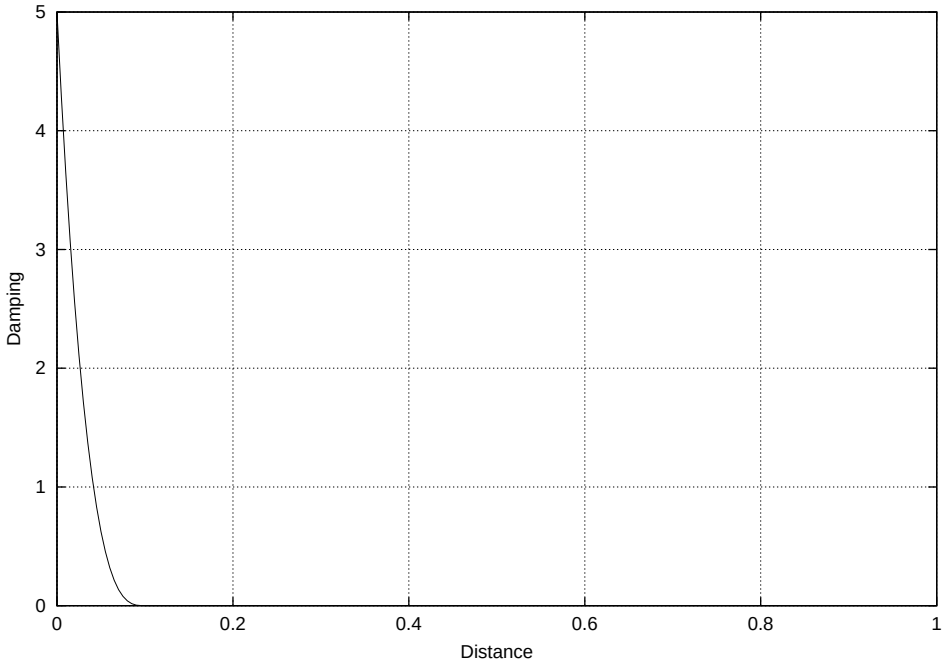

Figure 4.2: The damping function in one dimension

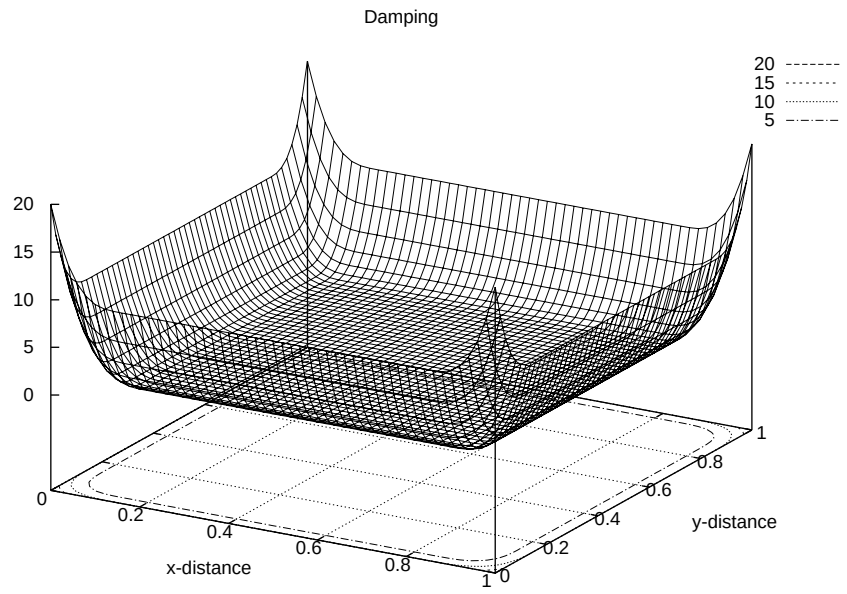

Figure 4.3: The damping function in two and three dimensions

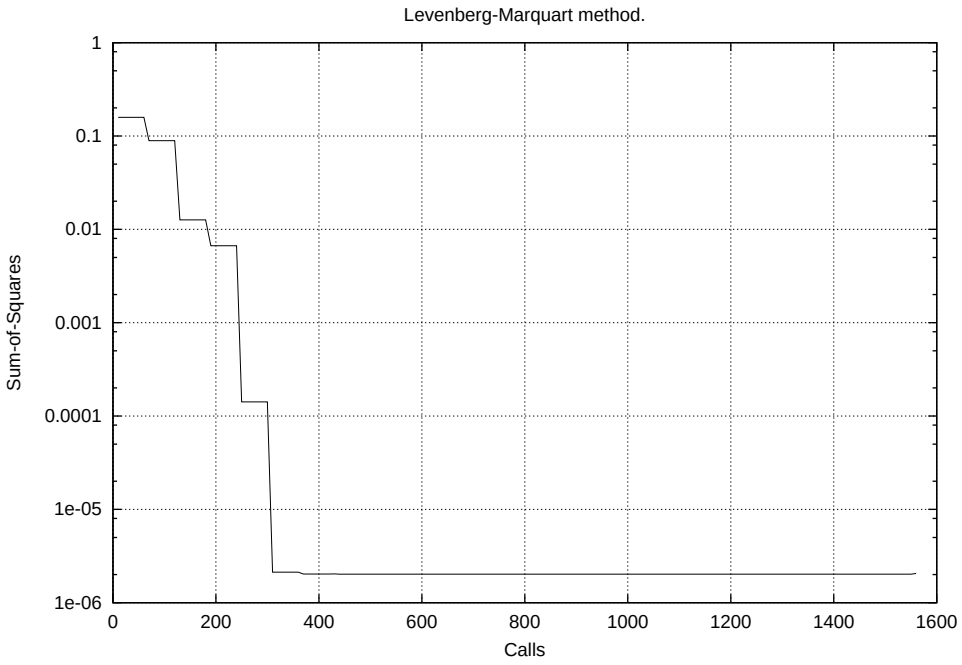

Figure 4.4: The plot of the Sum-of-Squares objective function as used by the Levenberg-Marquardt method.

and a new plateau then forms there. When one increases the parameter space dimension N, the longer the optimisation would take to converge.

As you can see in figure 4.4 the minimisation technique converges fast and, when it reaches its supposed minimum, it stays at that value for a while longer before stopping. That seems to indicate the Gauss-Newton behaviour is dominant in the LM. One drawback of this is that, when the objective function is near its minimum, the input parameters would also freeze into place with their latest values. If these values are such that the material property function has not been recovered properly, then the input parameter vector has stopped at the wrong place, even though the objective function (sum-of-squares) minimised properly. Many attempts (by trial-and-error) were made by changing the MERLIN input parameter starting vector and even altering or tweaking the model to fit the data properly. No penalty term has yet been introduced to the objective function to regularise it.

4.1.3 Use of units

Our theoretical framework can be used for any metric unit system (e.g. the cgs, mks (SI) or British units) since they involve only space, time and mass units. We do not deal explicitly with mass, rather it is contained implicitly in the material properties of the mass density and bulk modulus. The wave function can be unitless or in the case of displacements in the earth's crust have a length unit.

However, our physical problem is purely theoretical with only mathematical computations, and no real data from the field or any sources are tested. It is also simplified in the sense that space in each dimension (and possibly time also) runs from zero to unity. Mass density is unity throughout and does not feature in the calculations. The bulk modulus is between one and four for both the interface and gaussian feature. The wave function has unity amplitude initially until it interacts with the interface or feature. We concern ourselves only with the basic setup and therefore it does not matter which metric system of units one uses.

It is only when one starts testing this basic theory with the real data supported by a survey, that measurement units will become useful. Then we can say that we work for instance with a distance of 6 kilometers or an area of 20 square miles. But since the theory has not been tested, the values of all quantities presented remain dimensionless.

4.2 One-dimensional Calculations

4.2.1 One interface

Wave Propagation

We need to generate a wave amplitude $u(x,t)$ that satisfies the following differential equation (in one dimension):

$$\rho(x)\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial}{\partial x}\left(\lambda(x)\frac{\partial u(x,t)}{\partial x}\right) = \frac{\partial \lambda(x)}{\partial x}\frac{\partial u(x,t)}{\partial x} + \lambda(x)\frac{\partial^2 u(x,t)}{\partial x^2}, \qquad (4.1)$$

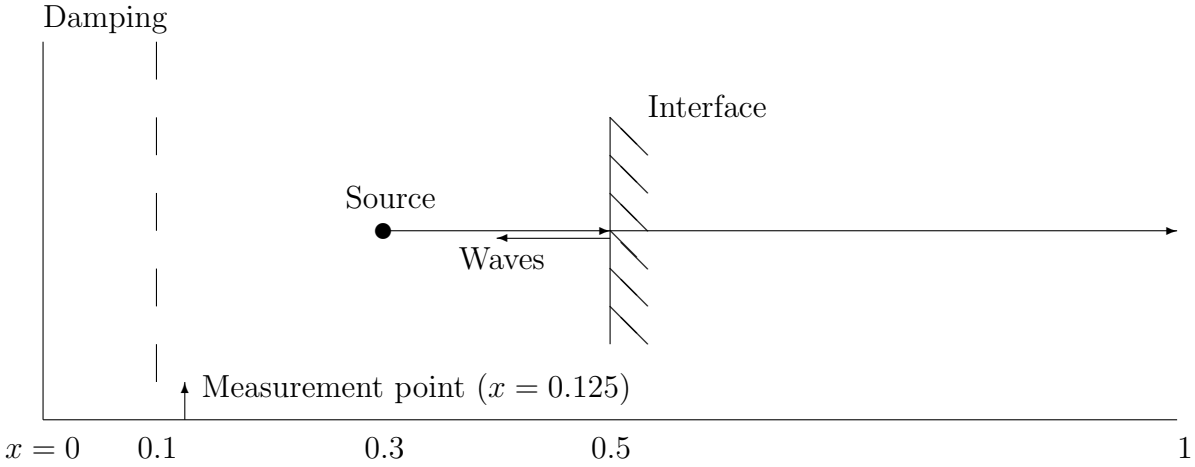

Figure 4.5: The physical setup for wave propagation in a one-dimensional medium with one interface

where $\rho(x)$ and $\lambda(x)$ are the mass density and first Lamé parameter of the material which can be obtained for $\mu(x) \equiv 0$ (second Lamé parameter) and $u = u(x,t)$ from Eq.2.30. The solution to the wave equation is composed of an incoming and an outgoing wave:

$$u(x,t) = u_-(x+ct) + u_+(x-ct). \tag{4.2}$$

For the moment we consider only the outgoing wave $u_+(x-ct)$. This is not a boundary value problem, but one with initial conditions only. The functional form for this is Gaussian:

$$u(x,0) = \exp(-\alpha(x-a)^2),$$

$$u(x,\Delta t) = \exp(-\alpha(x-a-\Delta t)^2),$$

where Δt is the time-step, $\alpha = 200^2 = 4 \times 10^4$ is the Gaussian sharpness and $a = 0.3$ is its initial position.

The space and time, in which the amplitude moves, is discretised into a number of equal steps in either dimension. The total distance is $d = h \cdot N = 1$ with space-step $h = 0.005$ and numerical grid density $N = 200$. The total time is $t = dt \cdot M = 1$ with time-step $dt = 0.002$ and number of time steps $M = 1000$. The Courant-Friedrichs-Lewy (CFL) condition for this differential equation is satisfied: $c\,dt/h = 2c/5 < 1$ where $c = \sqrt{\mu(x)/\rho}$ (see below) is between 1 and 2.

Figure 4.6 shows the wave amplitude over both space and time. In this plot the amplitude starts at $a = 0.3$ with time reset at zero. After 0.2 time-units it encounters the single interface, part of which reflects back to the original boundary and part transmits to the other side. One can see that the reflected wave is absorbed on $t = 0.7$ at the original boundary with the damping function as discussed in section 4.1.1. The transmitted wave travels at twice the velocity $c(x) = \sqrt{\lambda(x)/\rho_o} = 2$ than before (see next subsection on the material properties), and reflects at the opposite boundary after a further 0.25 time-units ($t = 0.45$). A quarter of a time-unit later a reverse transmission and inner reflection occurs at the interface ($t = 0.7$) as the boundary reflected wave comes back from below the interface. The inner reflections below the interface repeats the behaviour of boundary reflections and interface interactions from below, every half time-unit as the propagation time carries on. The reverse transmitted waves from below are also absorbed by the original boundary every half time-unit. Both transmission and reflection amplitude sizes are reduced by these interaction events at the interface, while the opposite boundary is an infinitely hard surface that does not allow any transmission there and the reflected amplitude size is not reduced. The amplitudes also undergoes a sign change upon all reflections, but not for all transmissions. These are in accordance with the rules that determine the coefficients for reflections and transmissions of waves. For clarity we have a cleaner three-dimensional plot in Figure 4.7 where the grid density is reduced and propagation time only goes to unity.

Calculations are done of the wave amplitudes over propagation time but sampled at $x_m = 0.125$ (Figure 4.8) together with their differences with respect to the target model reference wave (Figures 4.9 and 4.10) as used for the objective function. They do not look much like the full wave plots discussed in the previous paragraph, because the fixed sampling position x_x is just 0.025 space-units above the damping inner boundary w_x and 0.175 space-units below where the wave is due to start. Three minima can be seen at $t = 0.55$, $t = 1.09$, and $t = 1.58$, that correspond to the absorbed signals returning to the original boundary. The differences are very small, in the range of between -0.0006 and 0.0008. When the differences are squared, they

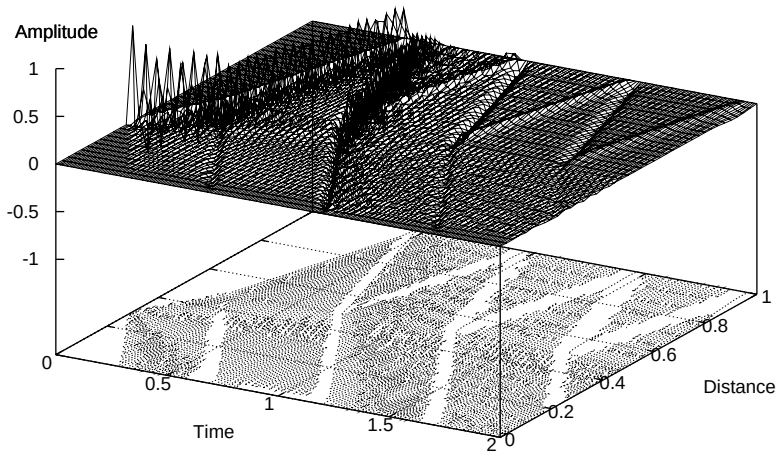

Figure 4.6: A plot of the wave propagation over the whole space-time grid for 1D single interface with damping only on the original side

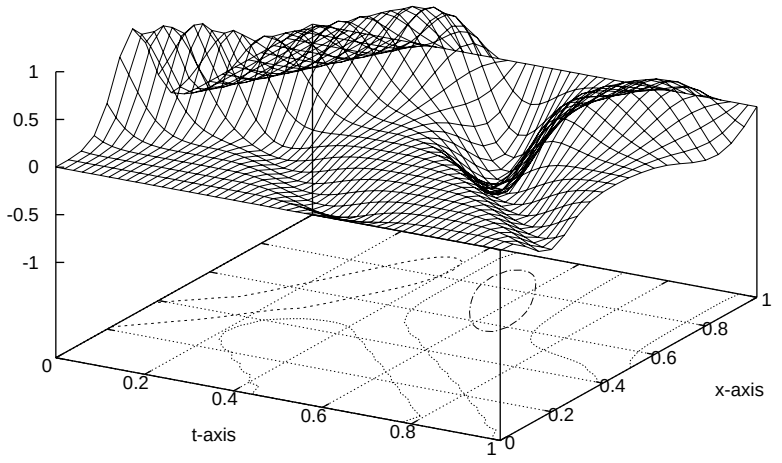

Figure 4.7: Same plot of the wave propagation but with $t \rightarrow 1$, not 2 for 1D single interface

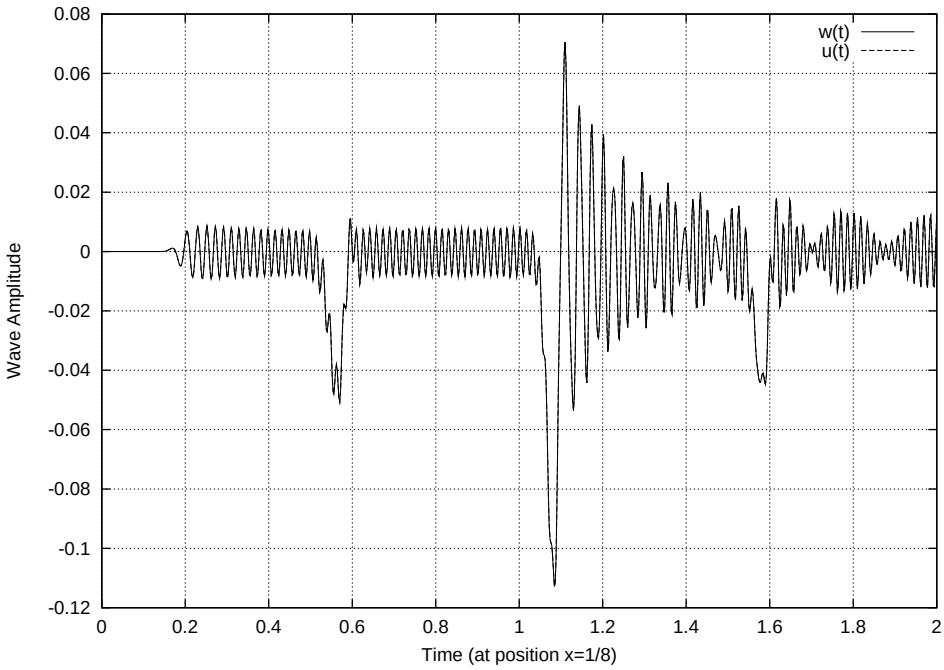

Figure 4.8: Plot of the optimised wave propagation at $x_{fixed} = 1/8$ for 1D single interface

show absolute values of up to 4.5×10^{-7}. They are most prominent at time ranges of $t\epsilon[0.55, 0.67]$, $[0.85, 1.13]$ and $[1.6, 2.0]$.

Material Property

Here the mass density and Lamé parameter related to the velocity of longitudinal waves are chosen as, respectively, $\rho(x) = \rho = 1$ and

$$\lambda(x) = \lambda_o + \frac{\lambda_f - \lambda_o}{1 + \exp(-\beta(x - b_x))},$$

where the typical values of the parameters are $\lambda_o = 1$ and $\lambda_f = 4$, with the transition made at position $b_x = 0.5$. The sharpness of this single layer interface function is $\beta = 200$ and is independent of h.

The number of minimisation input parameters is 55. This number is chosen to give a reasonable fit for more accurate results. A lower number would give less

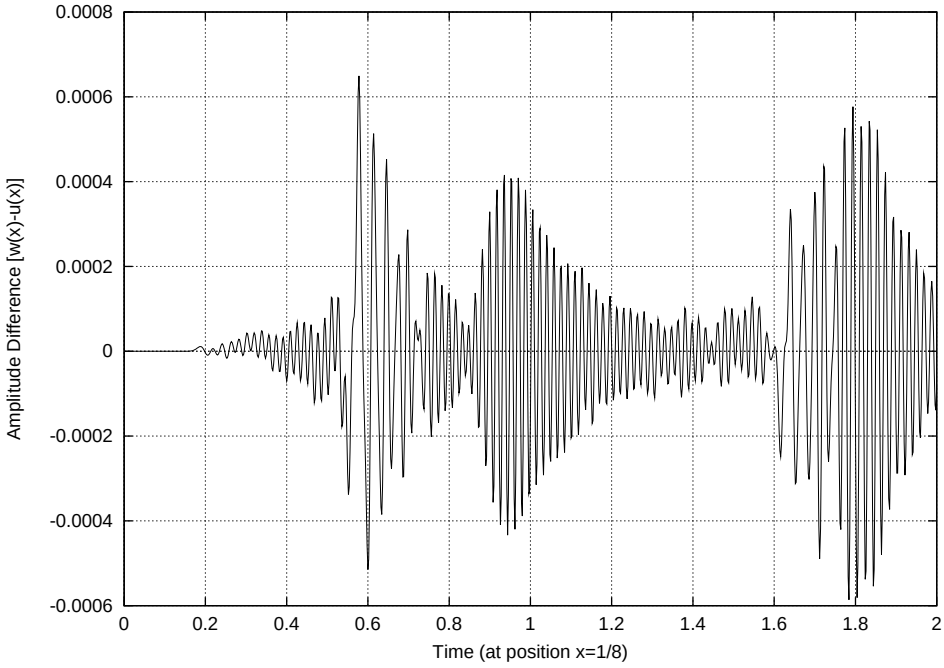

Figure 4.9: The final difference of waves at $x_m = 1/8$ for 1D single interface

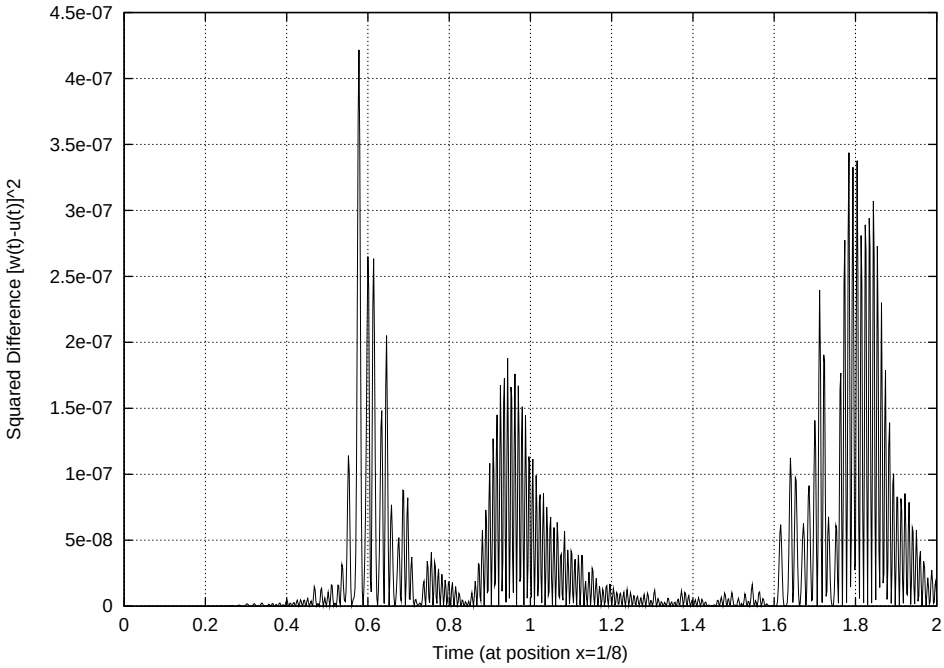

Figure 4.10: The squared difference of waves at $x_m = 1/8$ for 1D single interface

accurate results, but we have seen less accurate results at a higher numbers as well. Thus the number of parameters has to be chosen within an acceptable range. Not too low, but not too high either. These parameters are equated to the Fourier components. A Fourier sine series is used on the interval $x \in [b_o, b_f] = [0.3, 0.8]$. The equation for λ is thus

$$\lambda(x) = \lambda_o + (\lambda_f - \lambda_o) \left[\frac{x - b_o}{b_f - b_o} + \sum_{k=1}^{n} c_k \sin \left\{ k\pi \left(\frac{x - b_o}{b_f - b_o} \right) \right\} \right].$$

Define a function $\lambda_{FS}(x)$ by moving terms to the left hand side and keeping only the Fourier Series (FS) on the right hand side:

$$\lambda_{FS}(x) = \frac{\lambda(x) - \lambda_o}{\lambda_f - \lambda_o} - \frac{x - b_o}{b_f - b_o} = \sum_{k=1}^{n} c_k \sin \left[k\pi \left(\frac{x - b_o}{b_f - b_o} \right) \right],$$

where c_k are the Fourier components,

$$c_k = \frac{1}{\pi} \int \lambda_{FS}(x) \sin \left[k\pi X(x) \right] dx, \tag{4.3}$$

and $X(x) = (x - b_o)/(b_f - b_o)$.

MERLIN is initialised with zero components of the Fourier series as outlined here for one dimension on the interval given above. That is, it initialises to a piecewise linear function with fixed constant $\lambda_o = 1$ (the original material property) for $x \leq b_o$, arbitrary constant $\lambda_f = 4$ (the final material property) for $x \geq b_f$, and a straight line between b_o and b_f. The cosine terms are not used because at $x = b_o$ and $x = b_f$ they do not vanish.

Figures 4.11 and 4.12 show how close the minimisation has brought the material property to the target function. The first plot shows the single interface in the middle of the space and the final differences are almost indiscernible. The second plot shows the differences themselves and they have values to within ± 0.02. Dividing by the difference between the lower λ_o and upper λ_f levels $(\lambda_f - \lambda_o)$, the difference range is ± 0.00667, or plus-minus two thirds of a percent.

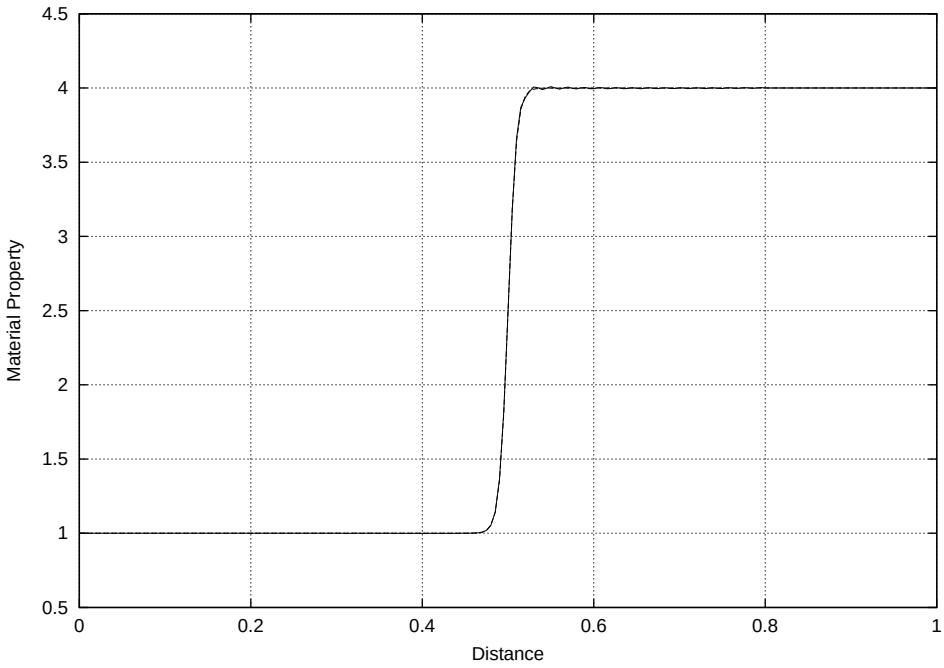

Figure 4.11: Plot of the optimised material property function λ for 1D single interface

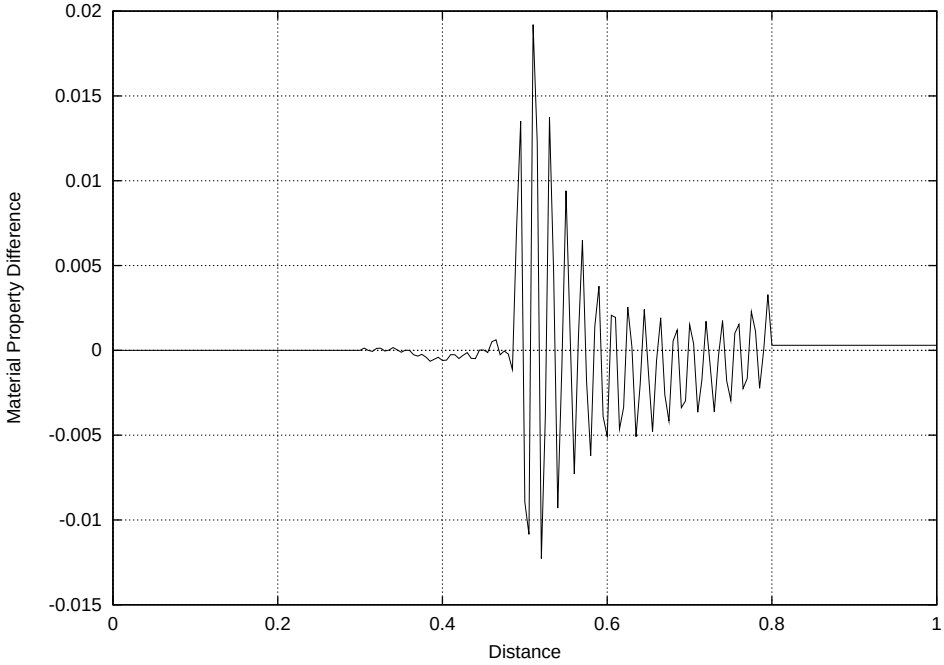

Figure 4.12: Plot of the optimised difference to its target function for 1D single interface

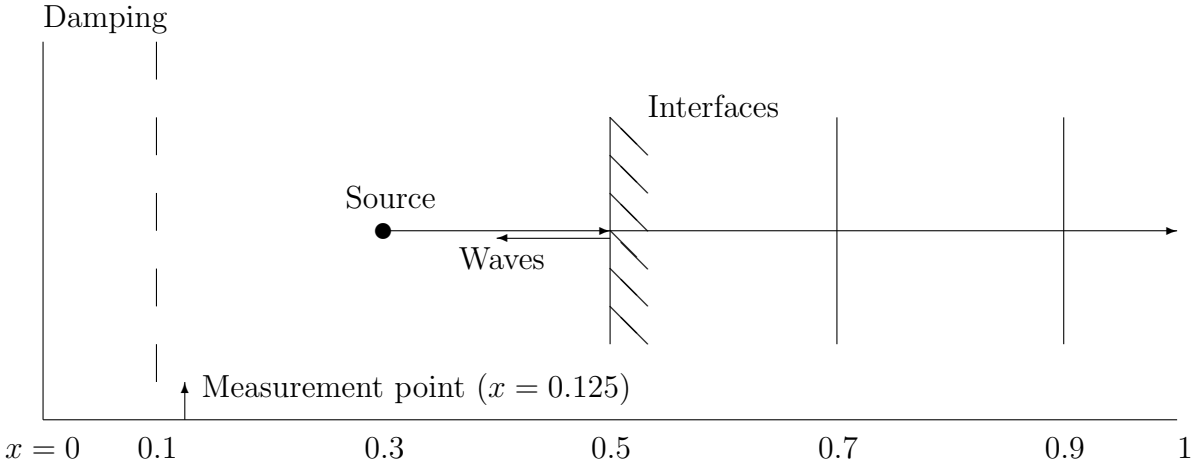

Figure 4.13: The physical setup for wave propagation in a one-dimensional medium with multiple interfaces

4.2.2 Multiple interfaces

Wave Propagation

We need to generate a wave amplitude $u(x,t)$ that satisfies the same differential equation (in one-dimension) as for just one interface in the previous section 4.2.1. Again this is no boundary value problem, but one with initial conditions only, with the same Gaussian functional form:

$$u(x,0) = \exp(-\alpha(x-a)^2), \qquad u(x,\Delta t) = \exp(-\alpha(x-a-\Delta t)^2),$$

where Δt is the time-step, $\alpha = 5 \cdot 200^2 = 2 \times 10^5$ is the Gaussian sharpness and $a = 0.3$ is its initial position.

The space and time are again discretised in the same way. The total distance is $d = h \cdot N = 1$ with space-step $h = 0.005$ and numerical grid density $N = 200$. The total time is $t = dt \cdot M = 1$ with time-step $dt = 0.002$ and number of time steps $M = 1000$. The Courant-Friedrichs-Lewy (CFL) condition for this differential equation is also satisfied: $c\,dt/h = 2c/5 < 1$ where $c = \sqrt{\mu(x)/\rho}$ (see below) is between 1 and 2.

Figure 4.14 shows again the wave amplitude over both space and time. One can see that all reflected waves returning to the original boundary are absorbed there. The discussion for the one dimensional single interface also applies here. But now

we have more than one interface which complicates the behaviour of wave propagation at later times. The plot is less clear after about propagation time $t = 1.2$, but reflections and some transmissions can still be seen. There is more overlap without interference of split waves, as they come together at one point at the same time and depart again unhindered on their original ways. All the interfaces still allow wave transmission and reflection with amplitudes reduced at each event, while the opposite boundary still allows only reflections and no transmission with no reduction in the amplitude.

Calculations are done of the wave amplitudes over propagation time but sampled at $x_m = 0.125$ (Figure 4.15) together with their differences with respect to the target model reference wave (Figures 4.16 and 4.17) as used for the objective function. Once again the plots are less clear after time $t = 1.2$ due to the more complicated nature of waves in multiple interfaces later on. Three minima are still visible before $t = 1.2$, which correspond to the absorbed signals returning to the original boundary, with the new minimum at about $t = 0.85$. This can only come from reflections at the new additional interface farther away from the original boundary. The differences are very small once again, in the range of ± 0.0002. The squared forms show absolute values of up to 3.5×10^{-8}. They are most prominent at time ranges of $t\epsilon[0.50, 0.75]$; $[0.75, 1.15]$ and $[1.75, 2.00]$.

Material Property

Here the mass density and Lamé parameter related to the velocity of longitudinal waves are chosen as $\rho(x) = \rho = 1$ and

$$\lambda(x) = \lambda_0 + \sum_{i=1}^{3} \frac{\lambda_i - \lambda_{i-1}}{1 + \exp(-\beta(x - b_{x_i}))},$$

where the typical values of the parameters are $\lambda_i = i + 1$, $b_{x_i} = (0.3 + 0.2i)$. The sharpness of this λ multi-layered interface function is $\beta = 200$ and once again independent of h.

The number of minimisation parameters (equated to Fourier components) is

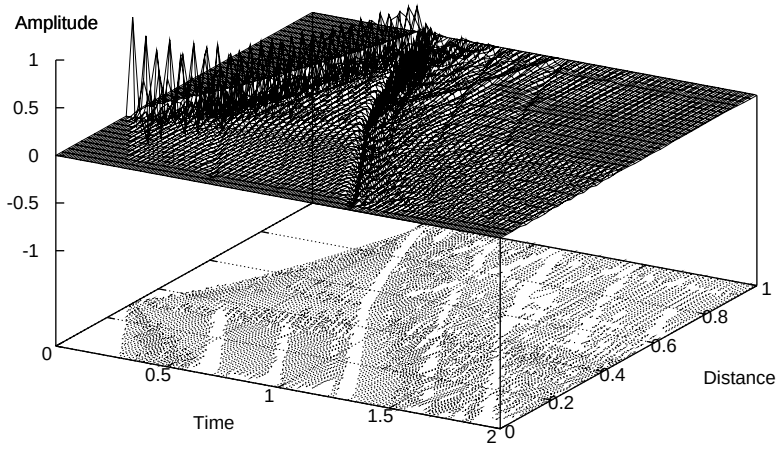

Figure 4.14: A plot of the wave propagation over the whole space-time grid with damping only on the original side for 1D multiple interfaces

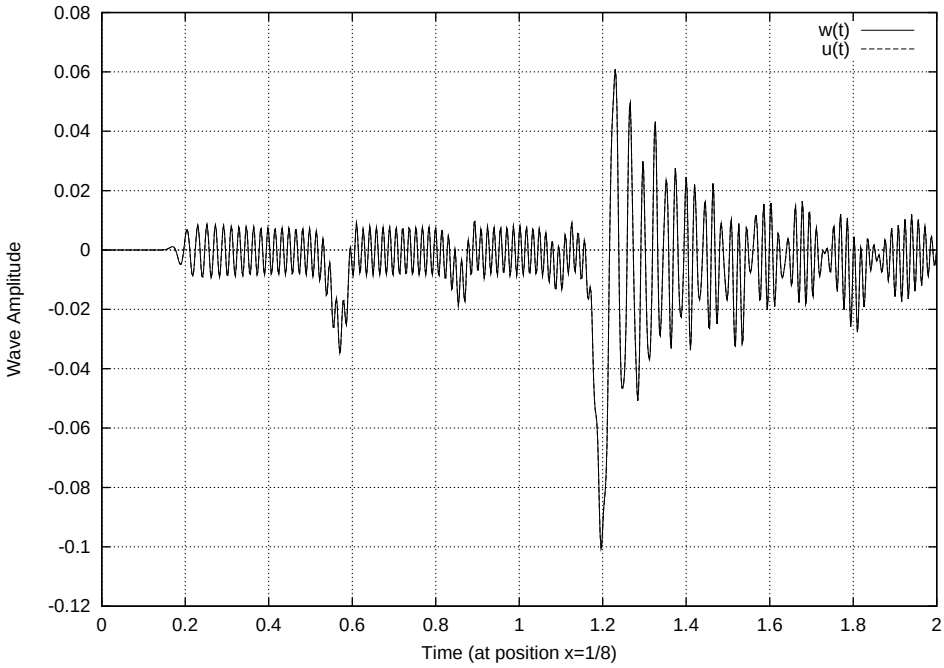

Figure 4.15: Plot of the optimised wave propagation at $x_{fixed} = 1/8$ for 1D multiple interfaces

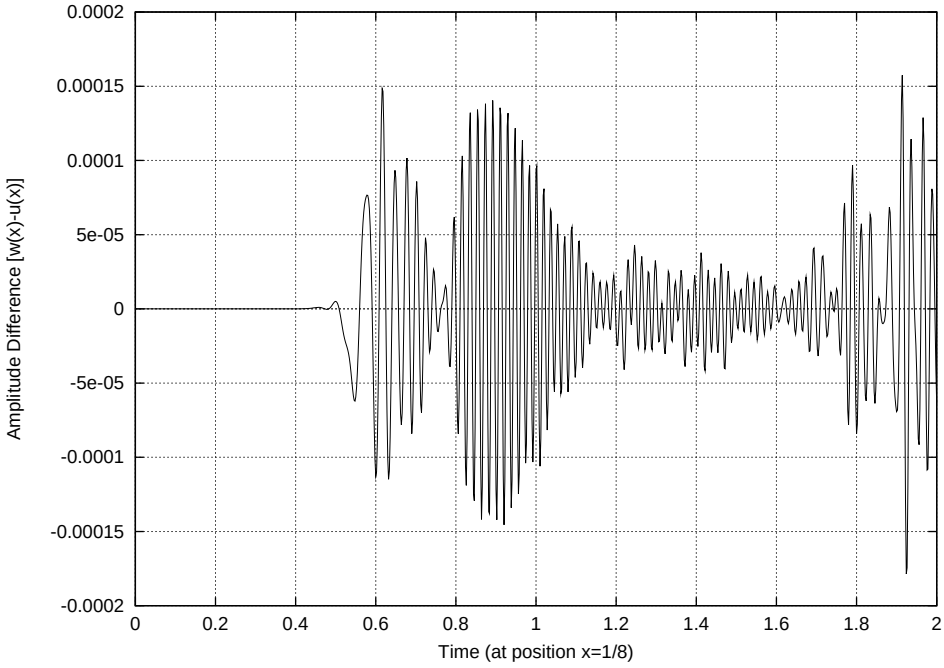

Figure 4.16: The final difference of waves at $x_m = 1/8$ for 1D multiple interfaces

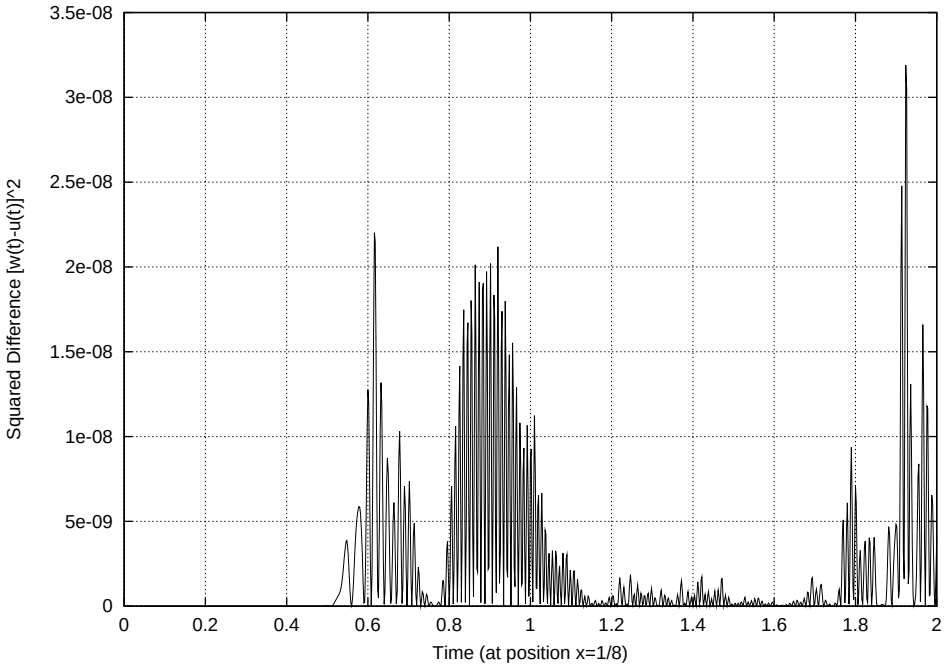

Figure 4.17: The squared difference of waves at $x_m = 1/8$ for 1D multiple interfaces

60. Same accuracy considerations apply with the results than in the previous case and two dimensional interface problems treated below. A Fourier sine series is used on the interval $x \in [b_o, b_f] = [0.45, 0.95]$. The equation for λ is thus

$$\lambda(x) = \lambda_o + (\lambda_f - \lambda_o) \left[X(x) + \sum_{k=1}^{n} c_k \sin \{k\pi X(x)\} \right].$$

Define again a function $\lambda_{FS}(x)$ where c_k are the Fourier components,

$$c_k = \frac{1}{\pi} \int \lambda_{FS}(x) \sin [k\pi X(x)] \, dx, \tag{4.4}$$

and $X(x) = (x - b_o)/(b_f - b_o)$. The λ_o and λ_f correspond to λ_i, $i = 0$ and $i = 3$ respectively. λ_f is also a Merlin input parameter for scaling the lambda function. MERLIN is initialised with zero components of the Fourier series.

Figures 4.18 and 4.19 show how close the minimisation has brought the material property to the target function. The first plot shows the multiple interfaces at positions $x = 0.5$, $x = 0.7$, and $x = 0.9$. Once again the final differences are almost indiscernible. The second plot shows the differences again and they have values to within ± 0.005 in the units of the material property. Dividing by the difference between the lower and upper levels $(\lambda_f - \lambda_o)$, the difference range is ± 0.001667, or plus-minus one sixth of a percent. This is four times better than one dimensional single interface results.

4.3 Two-dimensional Calculations

4.3.1 One interface

Wave Propagation

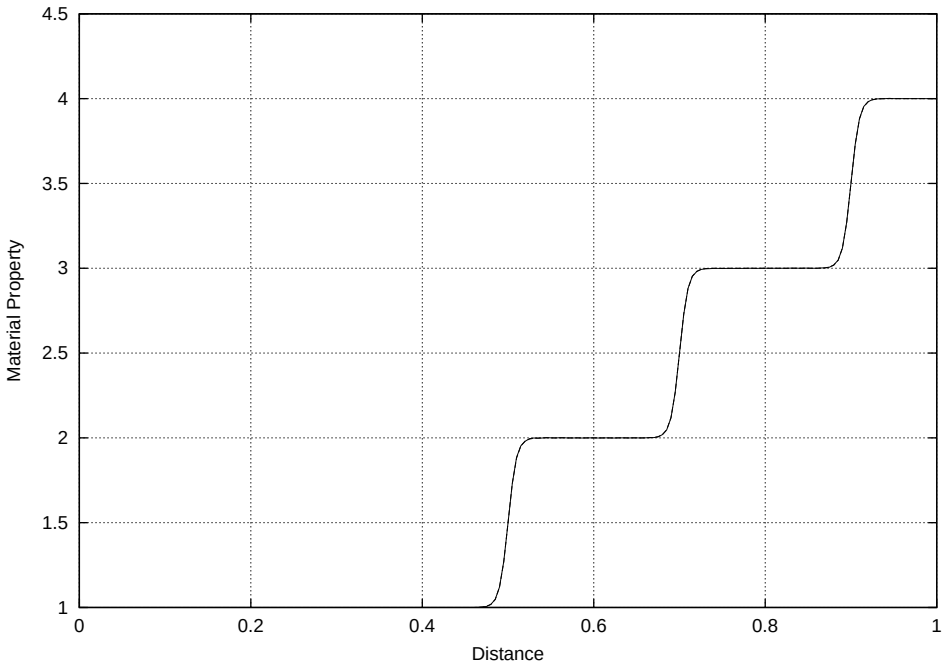

Figure 4.18: Plot of the optimised material property function λ for 1D multiple interfaces

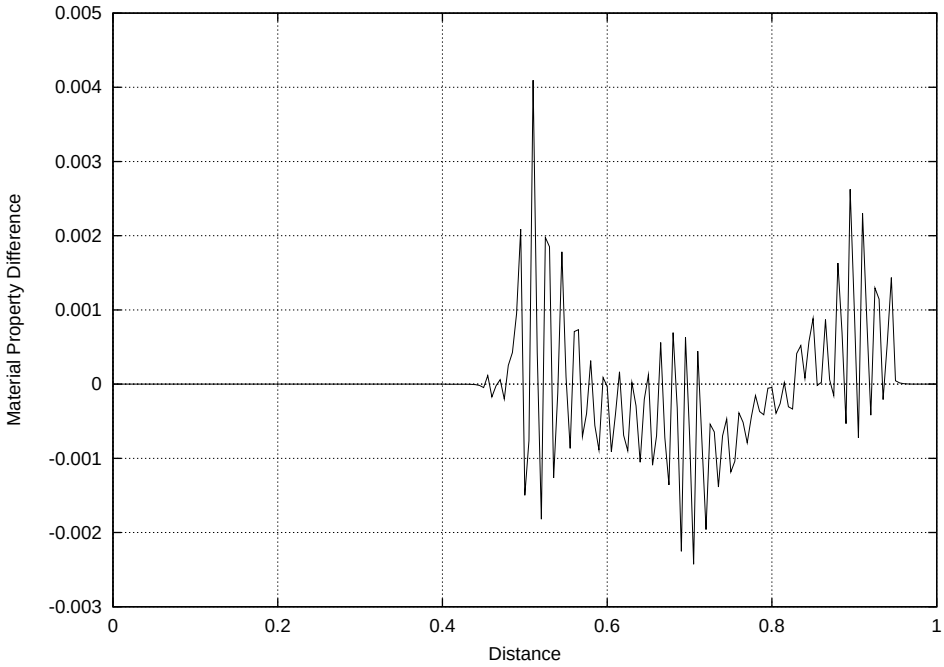

Figure 4.19: Plot of the optimised difference to its target function for 1D multiple interfaces

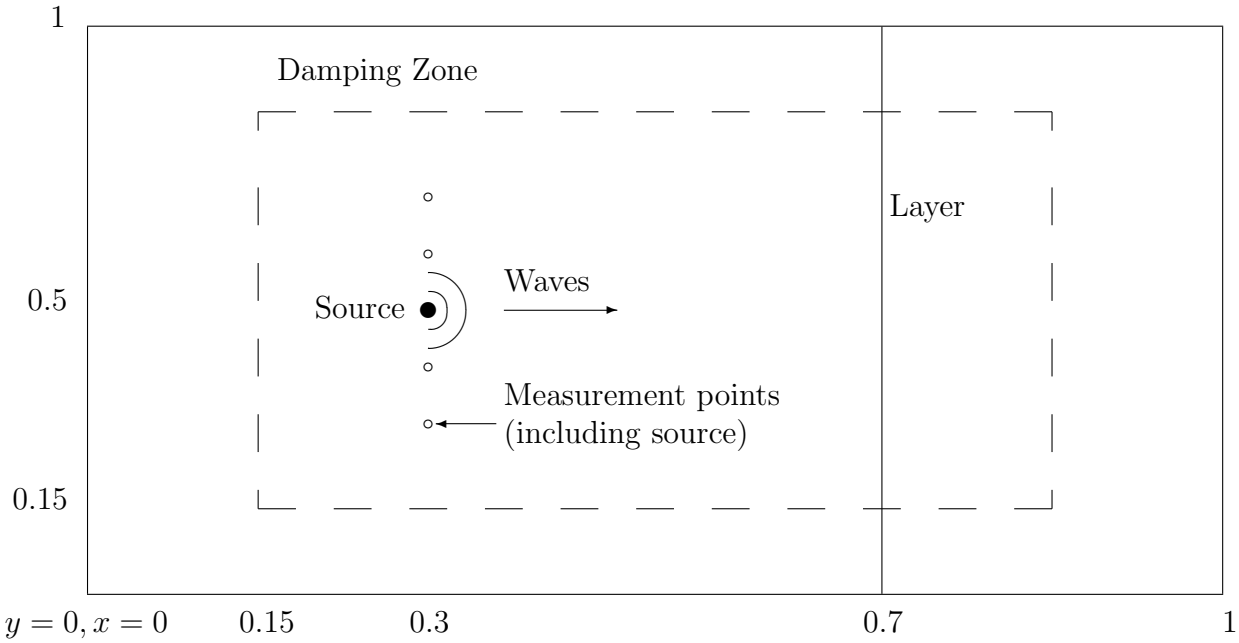

Figure 4.20: The physical setup for wave propagation in a medium with a layer in 2D.

We need to generate a wave amplitude $\mathbf{u}(\mathbf{r}, t) = \nabla\Phi(\mathbf{r}, t)$ with a potential $\Phi(\mathbf{r}, t)$ that satisfies the following differential equation in two dimensions. (See section 2.4):

$$\rho(\mathbf{r})\frac{\partial^2\Phi(\mathbf{r}, t)}{\partial t^2} = \lambda(\mathbf{r})\nabla^2\Phi(\mathbf{r}, t),$$

where ∇^2 is the Laplace operator and $\nabla = (\partial/\partial x, \partial/\partial y)$. Therefore,

$$\rho(x, y)\frac{\partial^2\Phi(x, y, t)}{\partial t^2} = \lambda(x, y)\left(\frac{\partial^2\Phi(x, y, t)}{\partial x^2} + \frac{\partial^2\Phi(x, y, t)}{\partial y^2}\right), \tag{4.5}$$

where $\rho(x, y)$ and $\lambda(x, y)$ are the mass density and first Lamé parameter of the material which can be obtained for $\mu(x, y) = 0$ (second Lamé parameter) and $\Phi = \Phi(x, y, t)$ from Eq.2.34 and $\mathbf{u} = \nabla\Phi$.

The radially expanding Gaussian wave from any point of origin is the initial amplitude for solution of the wave equation. That is

$$\Phi(\mathbf{r}, 0) = \exp(-\alpha r_A^2), \qquad \Phi(\mathbf{r}, dt) = \exp(-\alpha(r_A - dt)^2),$$

where $r_A = \sqrt{(x - a_x)^2 + (y - a_y)^2}$ is the wave radius, $\mathbf{a} = (a_x, a_y) = (0.3, 0.5)$ positions the point of origin, and the Gaussian sharpness is $\alpha = 5 \cdot 50^2 = 12500$. Here $h = 0.02$, $dt = 0.0025$ and $N = 50$, $M = 400$, all giving a total distance and

time of $d = Nh = 1$, $t = M\,dt = 1$. The CFL condition for this differential equation is satisfied: $c\,dt/h = 0.0025c/0.02 = c/8 < 1/\sqrt{2}$ where $c = \sqrt{\mu(\mathbf{r})/\rho}$ is between 1 and 2.

The boundary conditions for two dimensions now apply to the potential $\Phi(\mathbf{r})$. These boundary conditions are absorbing on all boundaries, thus $\Phi(\mathbf{r})$ and $\mathbf{u} = \nabla\Phi$ vanish on all boundaries. These apply to this test case and the other cases in two and three dimensions.

Multiple fixed position (or measurement points) are used for the sum-of-squares objective function. These are $r_i = (3, 3 + i)/10$, $i = 0, 1, 2, 3, 4$. The time development and final difference of the wave function at these points is shown in Fig.4.21 and 4.22. In the first plot these curves (but one) are flat in the beginning but then becomes oscillatory as the waves start to roll back and over these points. One curve is already wavy and has a much greater amplitude at the beginning than the other curves later on. The curves of the second plot are flat for up to about 0.4 time-units. No differences arise since the returned waves do not arrive at the measurement points before then. After the arrivals the differences are between 1×10^{-5} and -2.5×10^{-5}. When squared the absolute differences are up to 4×10^{-10}.

The full amplitude over the whole space is shown in Figure 4.23. This is actually a series of 9 time frames showing the propagation progress from a point on the one side of the x-axis and in the middle of the y-axis. A spherical outward leading wave expands from that point and one sees interference effects inside. In addition the waves are not perfect circles due to the coarseness of the space (fewer number of space-steps; only 50 by 50 steps). This can be improved, but with a price of too long computer runtimes. With damping on all boundaries, the wave and all interferences disappear. There is an interaction with the single interface at $x = 0.7$ and at the eighth frame. Due to the interference from behind, a reflection is barely visible. The transmitted wave speeds up to twice its velocity beyond the interface and has a greater wavelength. During all this time the amplitude is reduced from

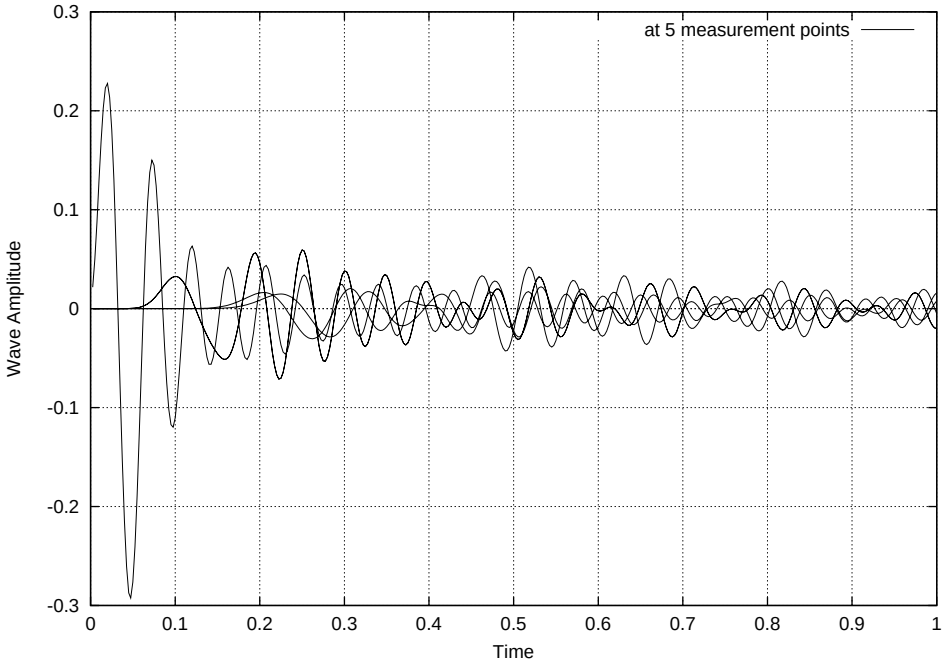

Figure 4.21: The wave amplitude in time at several measurement points for 2D single interface

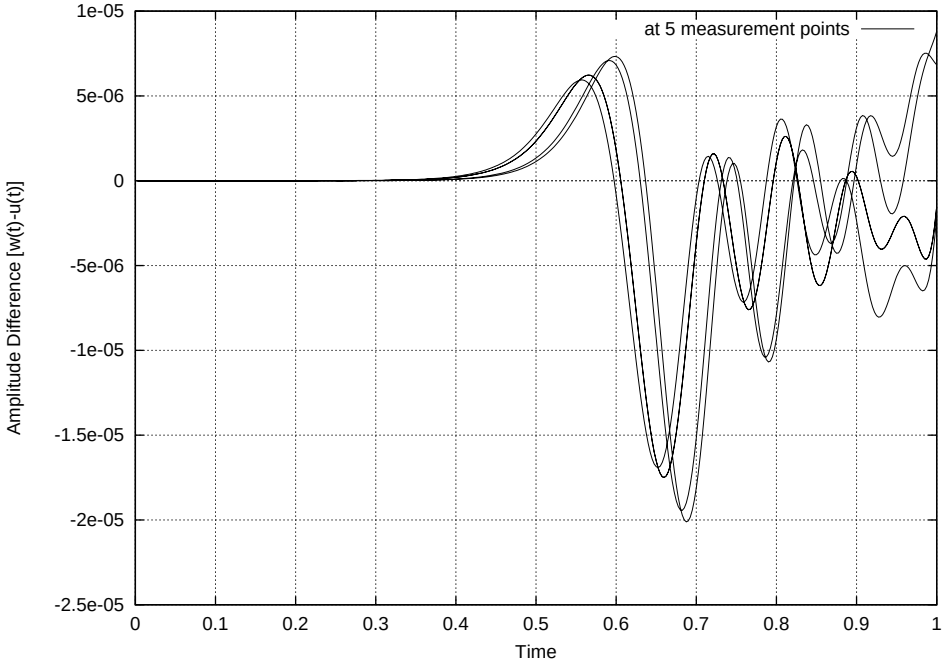

Figure 4.22: The final difference of amplitude at several measurement points for 2D single interface

unity at the start to 0.05 at the end.

Material Property

Here the λ is modelled by a single interface parallel to the y-axis [mass density $\rho(x, y) = 1$]:

$$\lambda(\mathbf{r}) = \lambda_o + \frac{\lambda_f - \lambda_o}{1 + \exp(-\beta(x - b_x))},$$

where the typical values of the parameters are the same again as in one dimension: $\lambda_o = 1$ and $\lambda_f = 4$, but the position was $b_x = 0.7$. Then $\beta = 50$ is the sharpness of λ.

The ansatz for the optimising function on the interval $x \in [b_o, b_f] = [0.5, 0.9]$, also independent of y in two dimensions, is

$$\lambda(\mathbf{r}) = \lambda_o + (\lambda_f - \lambda_o) \left[X(x) + \sum_{k=1}^{n} c_k \sin\left(k\pi X(x)\right) \right],$$

with the same defining function $\lambda_{FS}(\mathbf{r})$ within the Fourier coefficients c_k, and $X(x) = (x - b_o)/(b_f - b_o)$. Here the 40 Fourier components c_k are initialised to zero, giving the same piecewise initial function as in one dimension but now extended into and independent of the y-direction. Same accuracy considerations apply than in one-dimensional case to the chosen number of components.

The optimised material property function λ and its difference from the target is plotted in Fig.4.24 and 4.25. In the first plot the optimised function does not quite fit the target function. This is most pronounced with a dip at $x = 0.8$. Thus in the second plot the biggest negative difference at $x = 0.8$ touches at around -0.14, while the positive difference is only at around 0.04 in the units for the material property function.

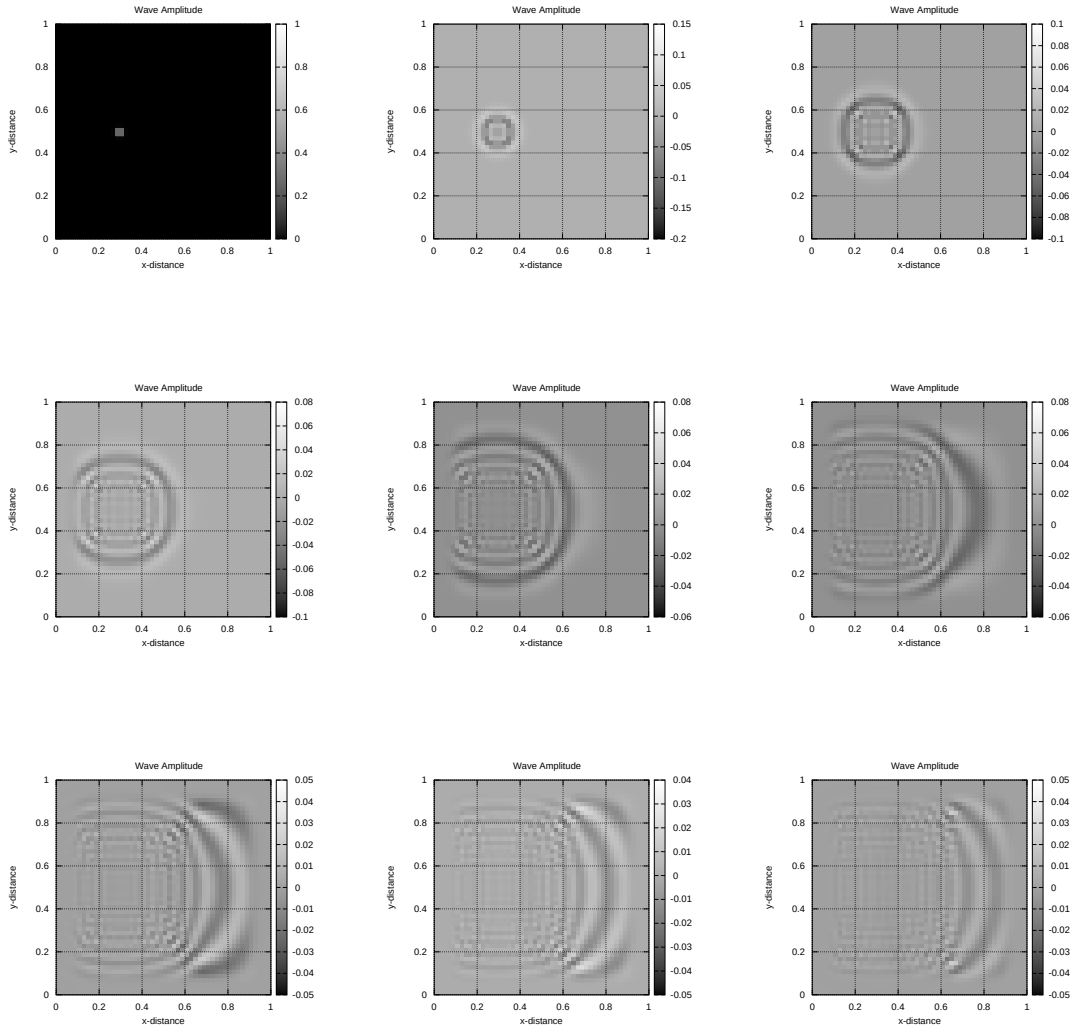

Figure 4.23: The amplitude in 2D space: time frames are from $t = 0$ to $t = 0.8$ in 0.10 seconds increments for 2D single interface

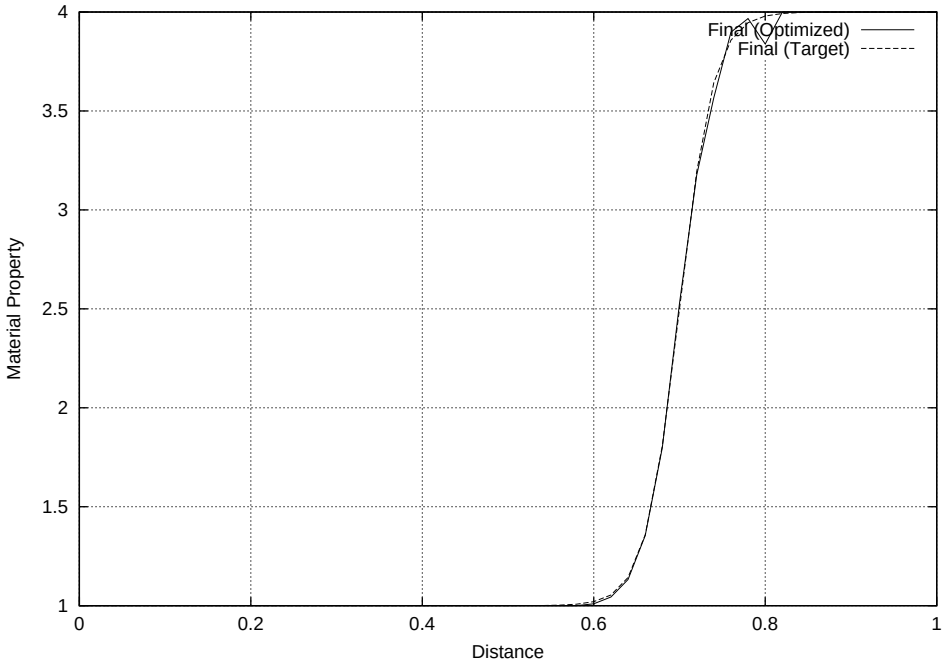

Figure 4.24: The optimised material property function λ for 2D single interface

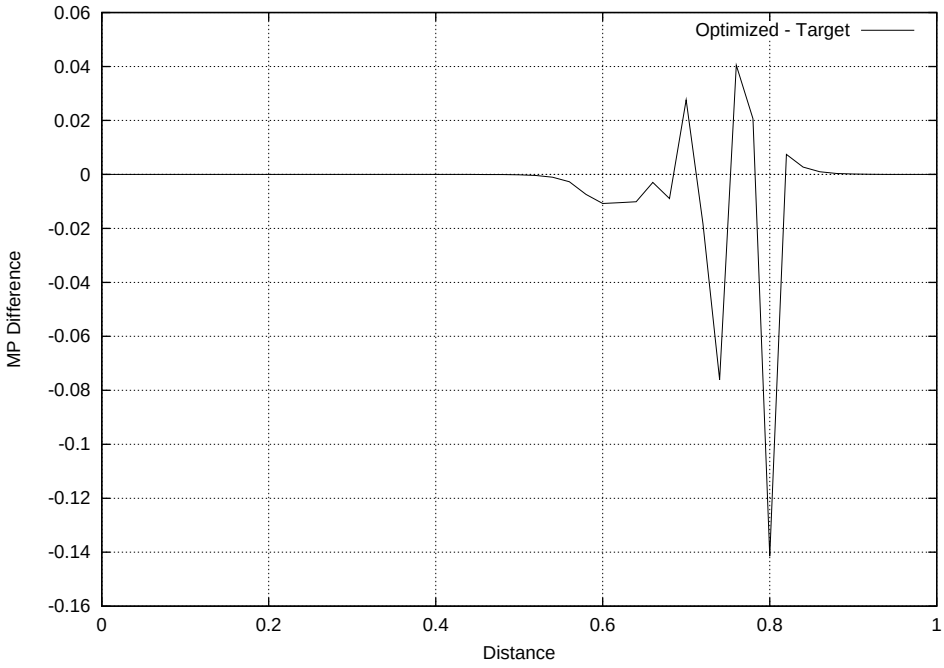

Figure 4.25: The optimised difference to its target function for 2D single interface

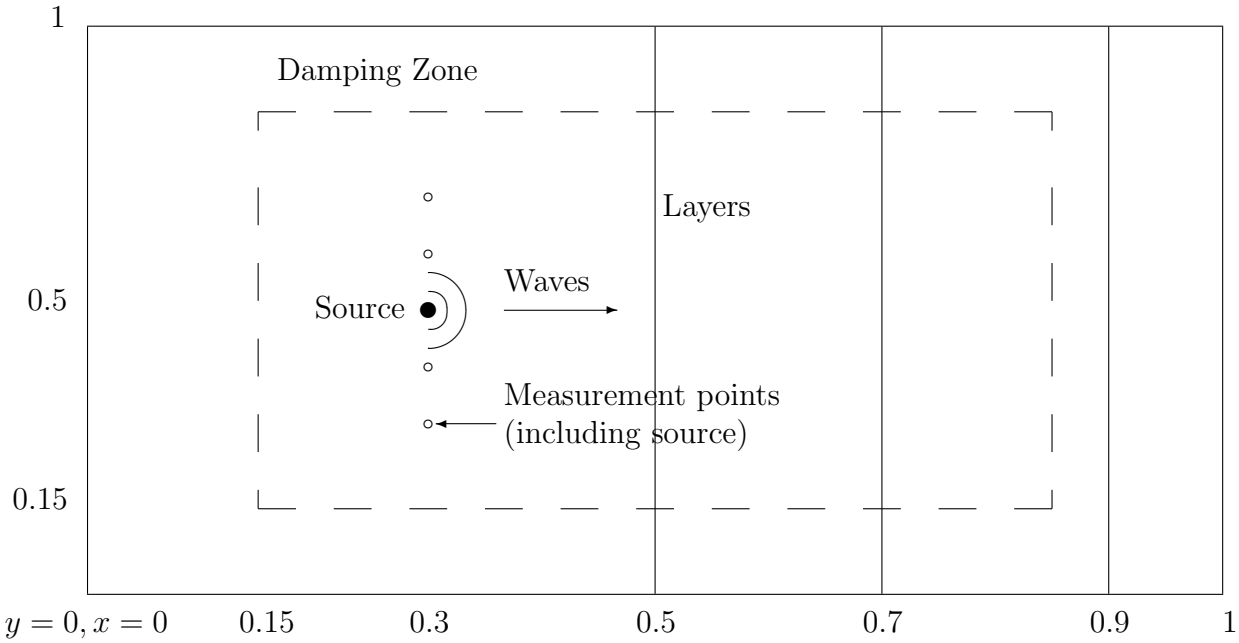

Figure 4.26: The physical setup for wave propagation in a medium with multiple layers in 2D.

4.3.2 Multiple interfaces

Wave Propagation

We need to generate a wave amplitude $\mathbf{u}(\mathbf{r}, t) = \nabla\Phi(\mathbf{r}, t)$ with a potential $\Phi(\mathbf{r}, t)$ that satisfies the same differential equation in two dimensions as in the previous section 4.3.1.

The radially expanding Gaussian wave from any point of origin is the initial amplitude for solution of the wave equation. That is

$$\Phi(\mathbf{r}, 0) = \exp(-\alpha r_A^2), \qquad \Phi(\mathbf{r}, dt) = \exp(-\alpha(r_A - dt)^2),$$

where $r_A = \sqrt{(x - a_x)^2 + (y - a_y)^2}$ is the wave radius, $\mathbf{a} = (a_x, a_y) = (0.3, 0.5)$ is the point of origin, and the Gaussian sharpness is $\alpha = 5 \cdot 40^2 = 8000$. Here $h = 0.025$, $dt = 0.0025$ and $N = 40$, $M = 800$ all giving a total distance and time of $d = Nh = 1$, $t = M\,dt = 2$. The CFL condition for this differential equation is satisfied: $c\,dt/h = 0.0025c/0.025 = c/100 < 1/\sqrt{2}$ where $c = \sqrt{\mu(\mathbf{r})/\rho}$ is between 1 and 2. The same absorbing boundary conditions as in section 4.3.1 apply here.

More than one fixed position (or measurement point) is used for the sum-of-

squares objective function. These are $r_i = (3, 3+i)/10$, $i = 0, 1, 2, 3, 4$. The time development and optimised difference of the wave function at these points is shown in Fig.4.27 and 4.28. The first plot shows basically similar behaviour as before. The curves of the second plot are flat for up to about 0.3 time-units. No differences arise since the reflected wave do not arrive at the measurement points before then. After the arrivals the differences are between 2.5×10^{-6} and -3×10^{-6}. When squared the absolute differences are up to 9×10^{-12}. This is a better result than that for the single interface.

The full amplitude over the whole space is shown in Figure 4.29. This is also a series of 9 time frames showing how the propagation progress from a point on the one side of the x-axis and in the middle of the y-axis. A spherical outward leading wave expands from that point and one again sees interference effects inside. In addition the waves are not perfect circles due to the coarseness of the space (only 40 by 40 space-steps). This can be improved, but with a price of too long computer runtimes. With damping on all boundaries, the wave and all interferences disappear. There is an interaction with the interfaces beginning at $x = 0.5$ and at the fifth frame a reflection is visible, despite the interferences from behind. The transmitted wave speeds up in steps, depending on the particular level, to twice its velocity beyond the last interface. In the last three frames standing waves seem to form after the first interface at position $x = 0.5$. During all this time the amplitude is reduced from unity at the start to 0.06 at the end.

Material Property

Here λ is modelled by a multi-layered interface with the layers parallel to the y-axis [mass density $\rho(x, y) = 1$]:

$$\lambda(\mathbf{r}) = \lambda_0 + \sum_{i=1}^{3} \frac{\lambda_i - \lambda_{i-1}}{1 + \exp(-\beta(x - b_{x_i}))},$$

where the typical values of the parameters are the same again as in one dimension: $\lambda_i = i + 1$, $b_{x_i} = (0.3 + 0.2i)$. $\beta = 40$ is the sharpness of λ.

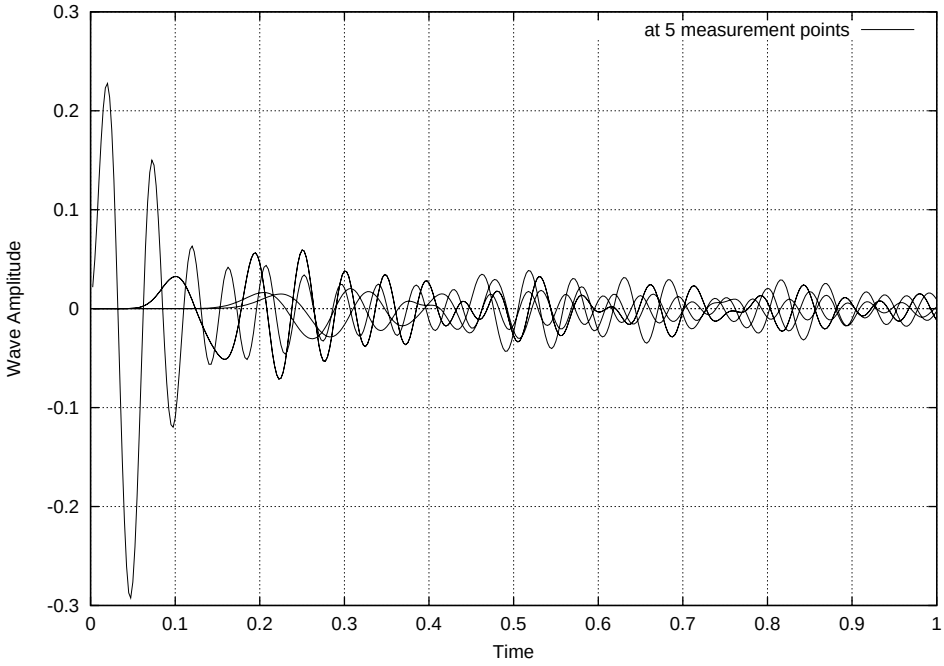

Figure 4.27: The wave amplitude in time at several measurement points for 2D multiple interfaces

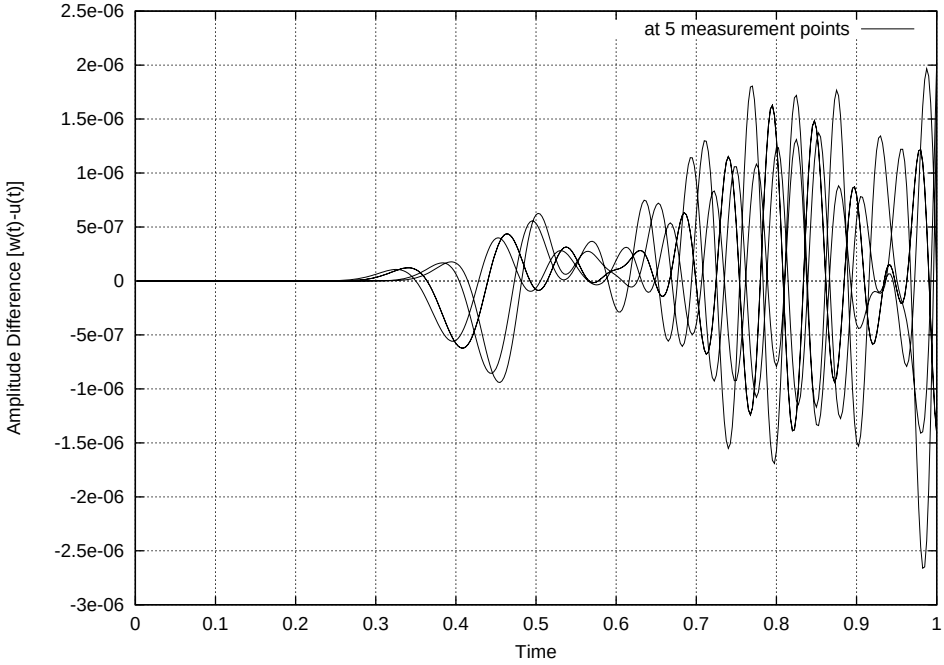

Figure 4.28: The final difference of amplitude at several measurement points for 2D multiple interfaces

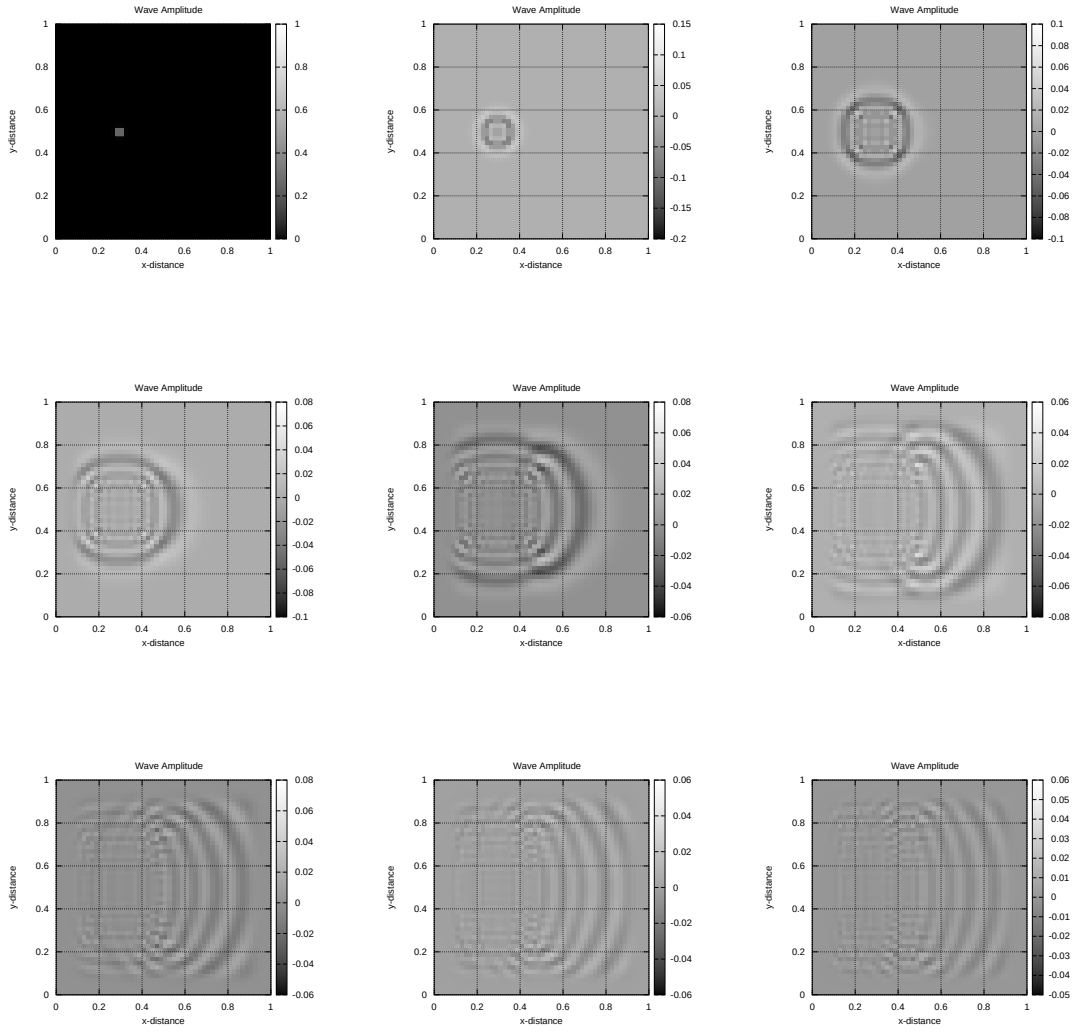

Figure 4.29: The amplitude in 2D space: time frames are from $t = 0$ to $t = 1.6$ in 0.20 seconds increments for 2D multiple interfaces

The optimising function on the interval $x \in [b_o, b_f] = [0.45, 0.95]$, also independent of y in two dimensions, is

$$\lambda(\mathbf{r}) = \lambda_o + (\lambda_f - \lambda_o) \left[X(x) + \sum_{k=1}^{n} c_k \sin\left(k\pi X(x)\right) \right],$$

with the same defining function $\lambda_{FS}(\mathbf{r})$ within the Fourier components c_k, and $X(x) = (x - b_o)/(b_f - b_o)$. The λ_o and λ_f again corresponds to λ_i, $i = 0$ and $i = 3$ respectively. λ_f is also a Merlin input parameter for scaling the lambda function. Here the 60 Fourier components c_k are initialised to zero, giving the same piecewise initial function as in one-dimension but now extended into and independent of the y-direction. Same accuracy considerations apply as well.

The optimised material property function λ and its difference from the target is plotted in Fig.4.30 and 4.31. In the first plot the optimised function is seen to differ at the third plateau from the target material property function. This may be due to the third interface positioned wholly inside the damping zone at the opposite side of the x-axis and may have little influence on returning waves at the measurement points. In the second plot the difference is between -0.025 and 0.03, in the units for the material property function. These differences arise because the Fourier method would optimise the components of the longest fundamental (resonant) wavelengths first, working its way through components to that of the shorter wavelengths until the objective function reaches a minimum. Once at the minimum, the rest of the components at the shortest end of wavelength spectrum simply freezes into non-optimised values. Then the associated short wavelengths are mismatched, leading to the oscillatory behaviour for the third and fourth plateaus in the material property function differences.

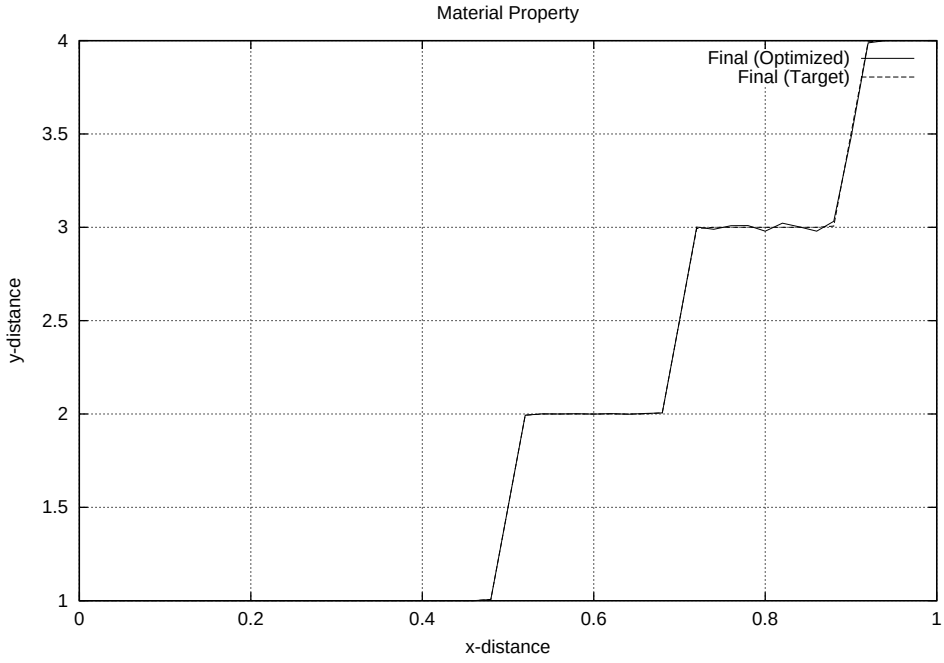

Figure 4.30: The optimised material property function λ for 2D multiple interfaces

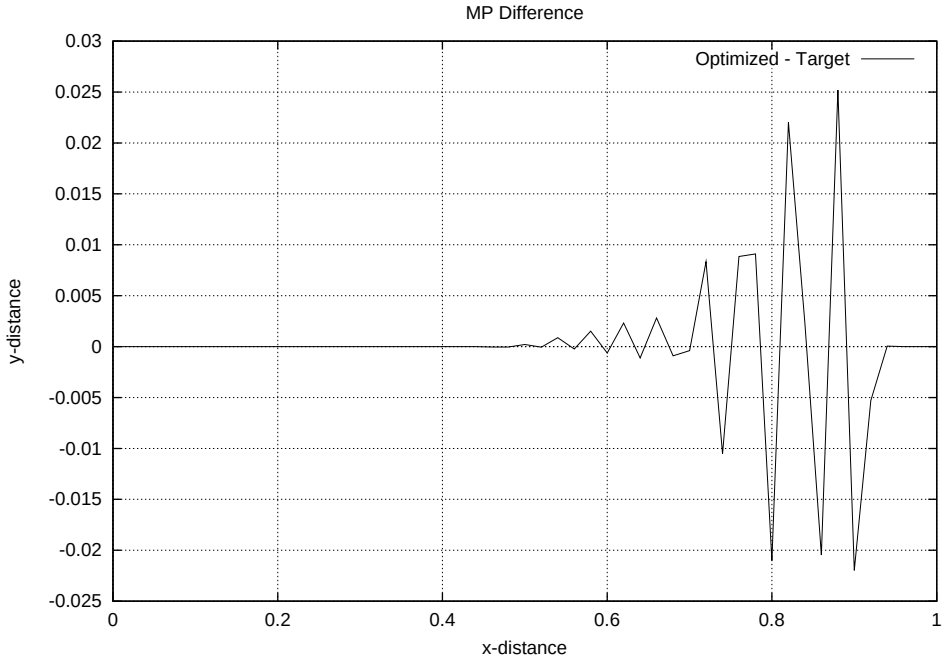

Figure 4.31: The optimised difference to its target function for 2D multiple interfaces

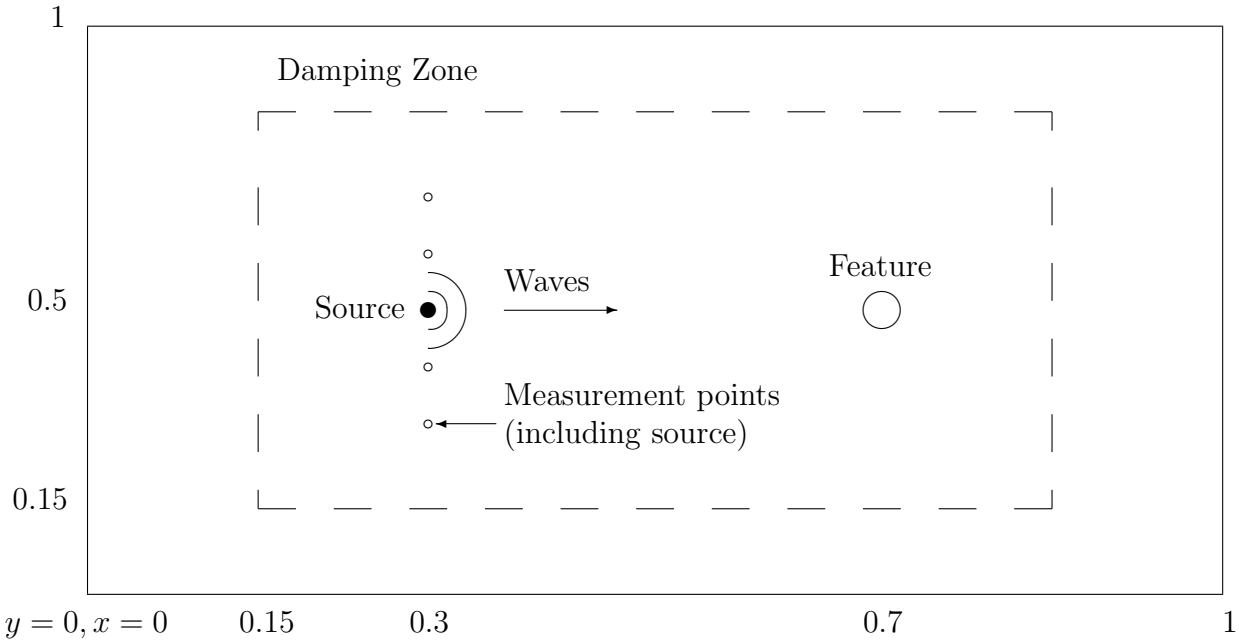

Figure 4.32: The physical setup for wave propagation in a medium with a feature in 2D.

4.3.3 Gaussian feature

Wave Propagation

The same differential wave equation in two dimensions for Φ applies here as in section 4.3.1.

The radially expanding Gaussian wave from any point of origin is the initial amplitude for solution of the wave equation. That is

$$\Phi(\mathbf{r}, 0) = \exp(-\alpha r^2), \qquad \Phi(\mathbf{r}, dt) = \exp(-\alpha(r - dt)^2),$$

where $r = \sqrt{(x - a_x)^2 + (y - a_y)^2}$ is the wave radius, $\mathbf{a} = (a_x, a_y) = (0.3, 0.5)$ positions the point of origin, and $\alpha = 1/5 \cdot 40^2 = 320$ is the Gaussian sharpness. Here too $h = 0.025$, $dt = 0.0025$ and $N = 40$ but $M = 400$. The CFL condition for this differential equation is satisfied: $c\,dt/h = 0.0025c/0.025 = c/10 < 1/\sqrt{2}$ where $c = \sqrt{\mu(\mathbf{r})/\rho}$ is between 1 and 2. The same absorbing boundary conditions as in section 4.3.1 apply here.

Once again, more than one fixed position (or measurement point) was used to improve the minimisation. These are $r_i = (3, 3 + i)/10$, $i = 0, 1, 2, 3, 4$. The time

development of the wave function and its difference from the target reference wave at these points are given in the diagram Figures 4.33 and 4.34. These peaks correspond to each of the five measurement points. In the first plot each peak position shows when in time the original wave arrived at each measurement point. The highest peak occurs at index $i = 2$ of r_i. There is a sort of pairing of the other four peaks. This means that the higher peak pair at around 0.1 time-units corresponds to the arrivals at $i = 1$ and $i = 3$, while the other pair at around 0.2 time-units are the arrivals at $i = 0$ and $i = 4$. Though barely discernable, reflections of the feature start to arrive at between $t = 0.6$ and $t = 0.8$. The second plot shows the curves are flat for up to 0.35 time-units. Then reflection effects starts to roll in with differences between -1×10^{-6} and 4×10^{-6}. The squared difference will be up to 1.6×10^{-11}.

The full wave function over the whole space follows in Figure 4.35 with 9 time development frames of wave propagation. Again the wave starts as a Gaussian peak at the point **a** on one side, which then dissolves into a spherically expanding wave. Interference behind the wave is less pronounced. Once again the waves are not perfect circles in regard of a fewer number of space-steps (40 by 40). This can be increased, but the computer runtimes will go into hours, if not days. When interacting with the feature (fifth frame), the wave speeds up to twice its velocity at peak position, then dissolves beyond the feature. Thus, a hole has formed in the wave around and beyond the feature. The feature does not create any interference of each part of the wave on either side. Each part just carries on expanding with the hole widening as time goes on. Damping is visible in the dark regions around the boundaries where the waves are absorbed. The amplitude reduces from unity to about 0.07 during this time.

Material Property

Here the λ is represented by a Gaussian feature, with the function [mass density $\rho(x, y)$ is unity over all space]:

$$\lambda(x, y) = \lambda_{env} + (\lambda_{peak} - \lambda_{env}) \exp(-\beta R^2),$$

where $\beta = 320$ is the Gaussian steepness, $R = \sqrt{(x - b_x)^2 + (y - b_y)^2}$ is the

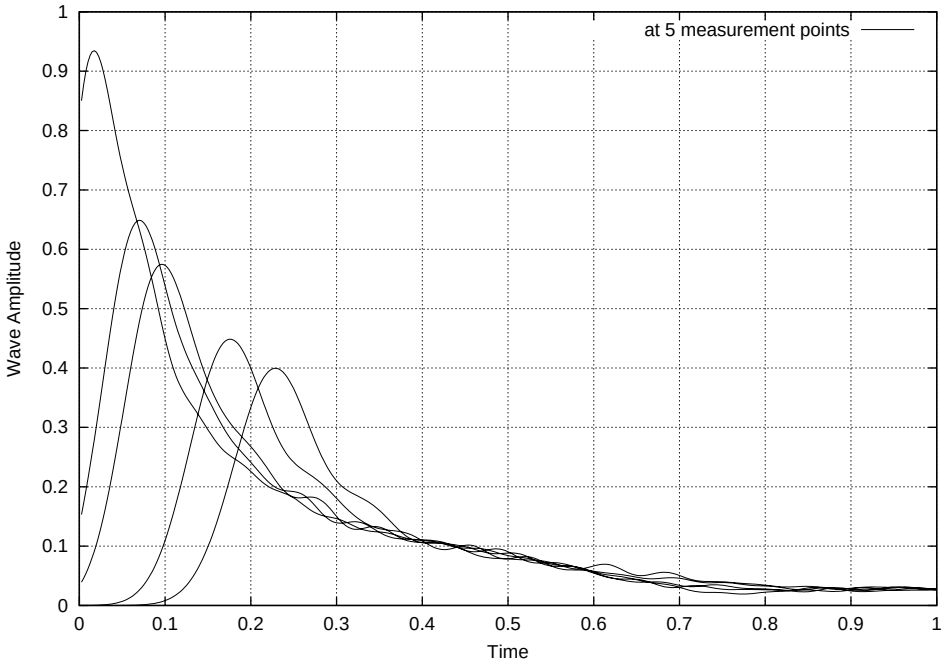

Figure 4.33: The amplitude of time at several measurement points for 2D Gaussian feature

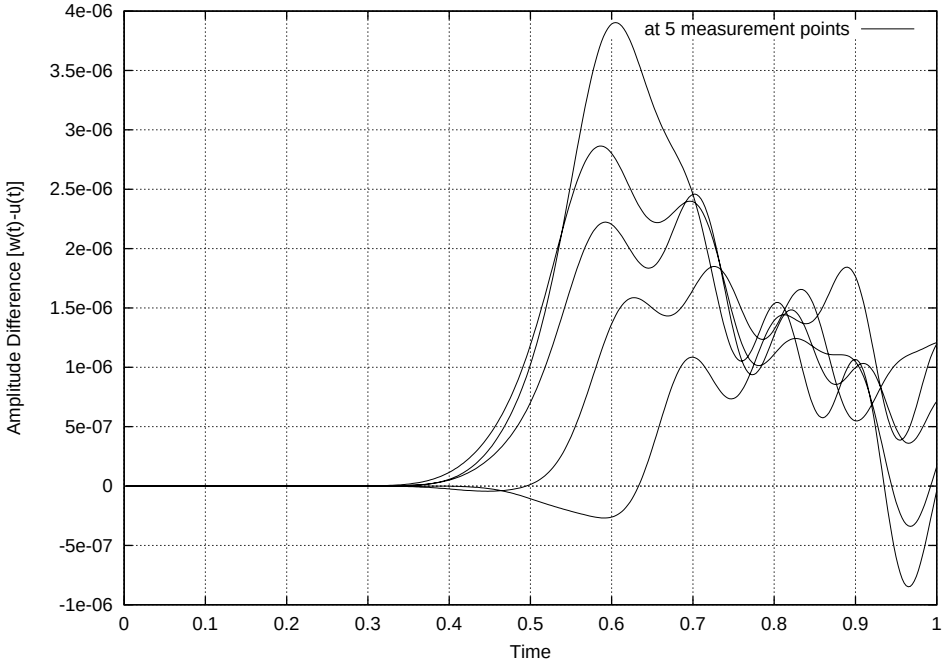

Figure 4.34: The amplitude difference at several measurement points for 2D Gaussian feature

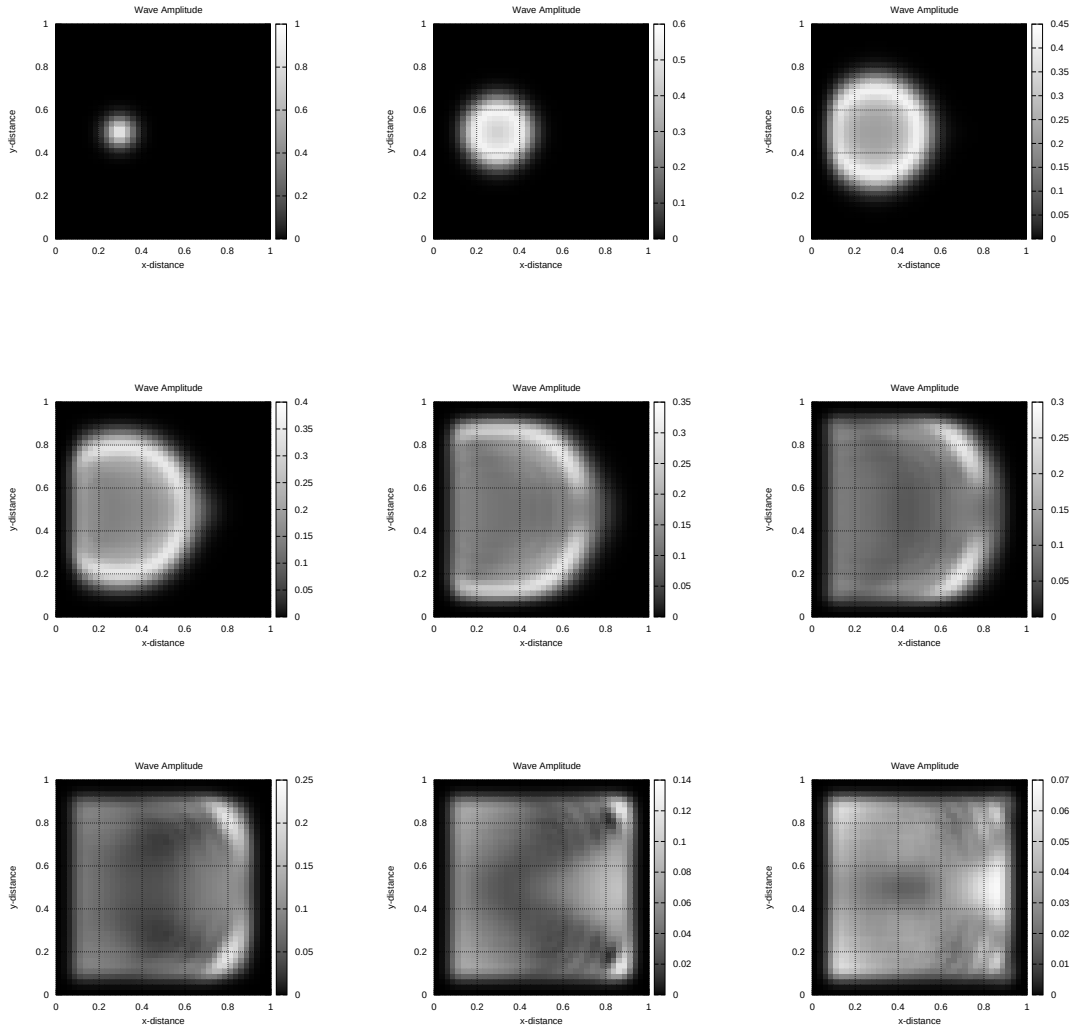

Figure 4.35: The amplitude in 2D space: time frames are from $t = 0$ to $t = 0.8$ in 0.10 seconds increments for 2D Gaussian feature

Gaussian radius, and $\mathbf{b} = (b_x, b_y) = (0.7, 0.5)$ is the position of the feature. λ_{env} is the environment level value and λ_{peak} is the value of the peak.

We use the finite-element method whose coefficients are now calculated from the λ-function above at the corner positions of the elements. The Laguerre polynomials are of fourth order, each operating on rectangular regions (the elements) of an area $H_x \times H_y$. Currently the side-lengths are $\mathbf{H} = (H_x, H_y) = (6, 6)h$ within each element, which makes them square. To accommodate the feature we have 4 squares, arranged in 2×2, whose common intersecting corners are at position $\mathbf{b}$, in the middle of the applied FEM-region.

The plots for two-dimensional problem follow. The optimised lambda function λ and its difference from its target is seen in Fig.4.36 and 4.37. The first plot shows a nice spherical Gaussian peak at position $\mathbf{b}$ on the other side of the space. In the current view of the second plot a negative difference is seen on the near side with two minima (the bigger one closer to the opposite y-boundary), and a positive difference on the far side with two maxima (again the bigger one closer to the opposite y-boundary). The differences range at ± 0.004.

4.4 Three-dimensional Calculations

4.4.1 Gaussian feature

Wave Propagation

We need to generate a wave amplitude $\mathbf{u}(\mathbf{r}, t) = \nabla\Phi(\mathbf{r}, t)$ for a potential Φ that satisfies the following differential equation in three dimension of Φ (see section 2.4):

$$\rho(\mathbf{r})\frac{\partial^2\Phi(\mathbf{r}, t)}{\partial t^2} = \lambda(\mathbf{r})\nabla^2\Phi(\mathbf{r}, t), \tag{4.6}$$

where ∇^2 is the Laplace operator and $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$. $\lambda(x, y, z)$ are the first Lamé parameter of the material which can be obtained for $\mu(x, y, z) = 0$ (second Lamé parameter) and $\Phi = \Phi(x, y, z, t)$ again from Eq.2.34.

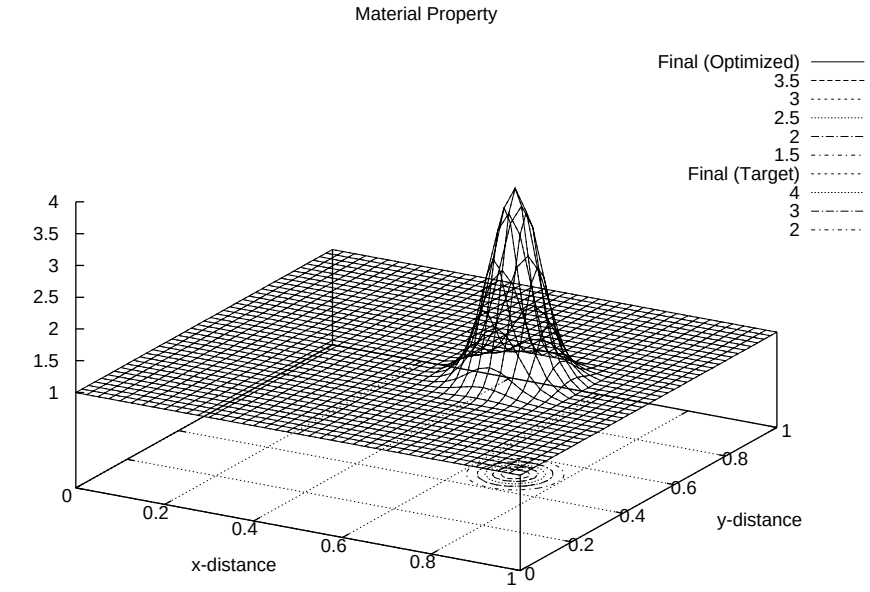

Figure 4.36: The optimised material property function λ for 2D Gaussian feature

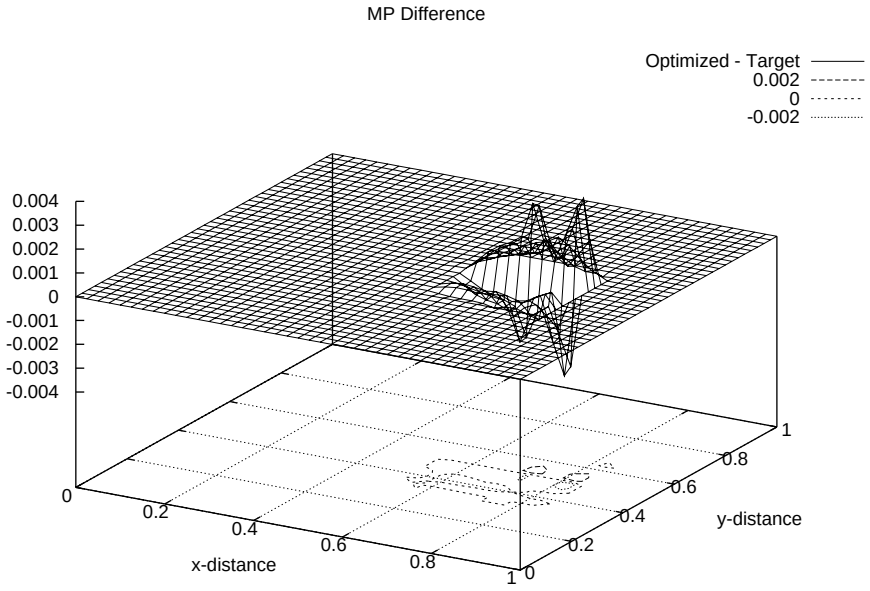

Figure 4.37: The optimised difference to its target function for 2D Gaussian feature

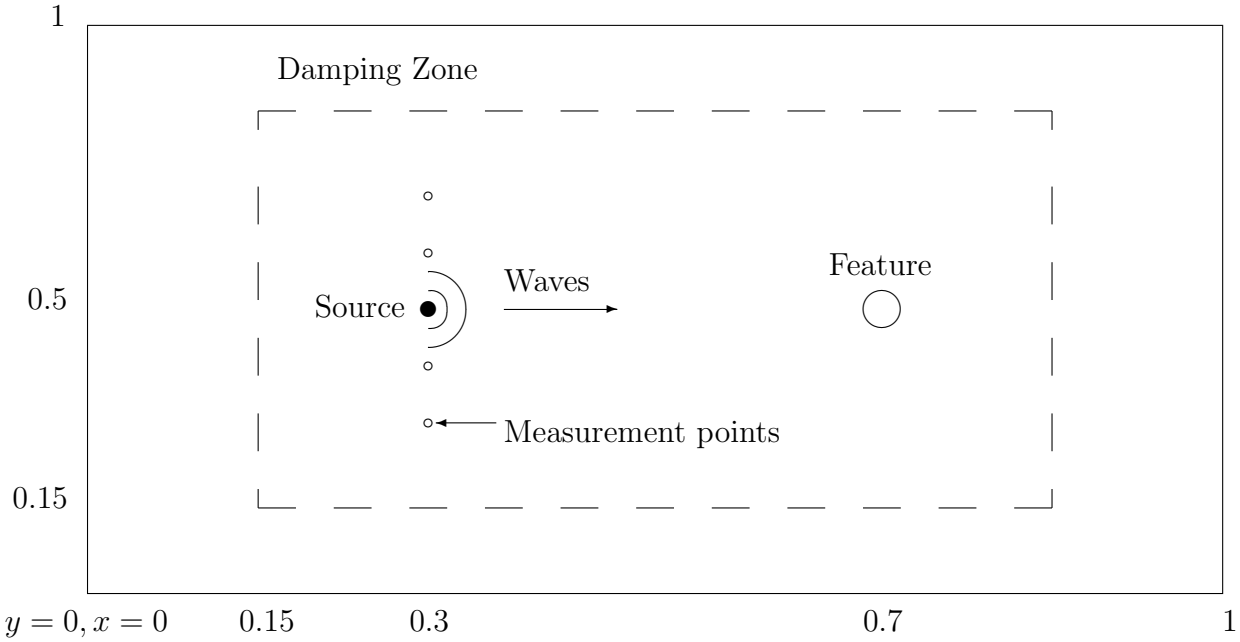

Figure 4.38: The physical setup for wave propagation in a medium with a feature in 3D.

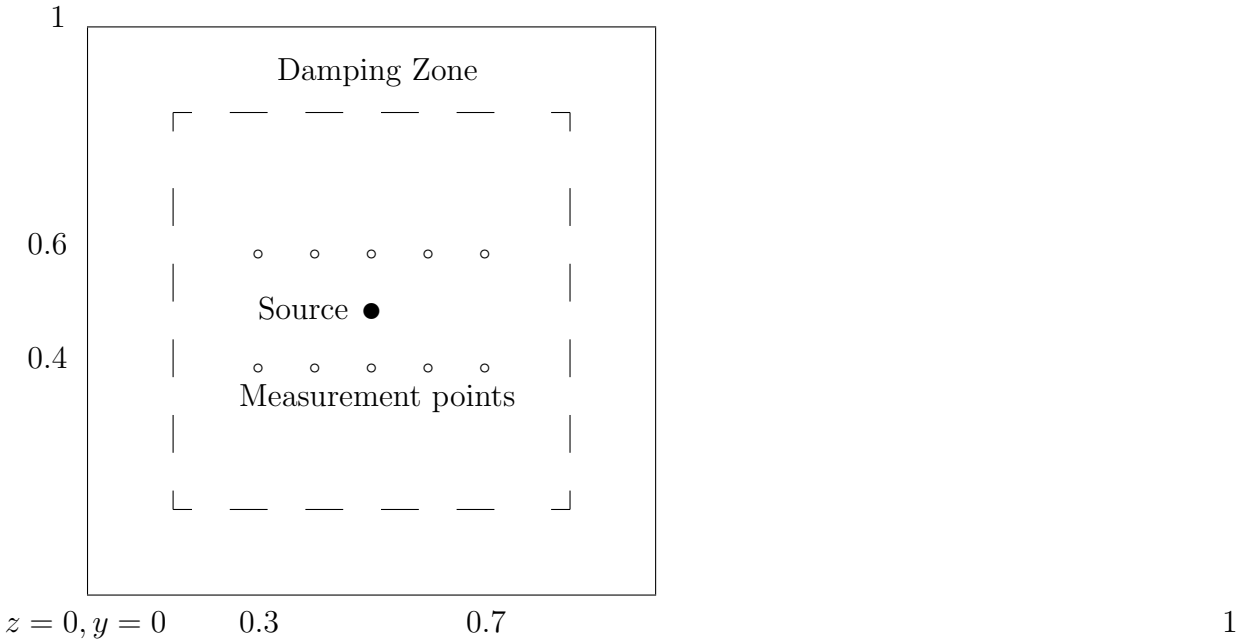

Figure 4.39: The same physical setup for wave propagation head-on in 3D

The spherically expanding Gaussian wave from any point of origin is the initial amplitude for solution of the wave equation. That is

$$\Phi(\mathbf{r}, 0) = \exp(-\alpha r^2), \qquad \Phi(\mathbf{r}, dt) = \exp(-\alpha(r - dt)^2),$$

where $r = \sqrt{(x - a_x)^2 + (y - a_y)^2 + (z - a_z)^2}$ is the wave radius, $\mathbf{a} = (a_x, a_y, a_z) = (0.3, 0.5, 0.5)$ positions the point of origin, and $\alpha = 1/5 \cdot 40^2 = 320$ is the Gaussian sharpness. Here too $h = 0.025$, $dt = 0.0025$ and $N = 40$ but $M = 400$. The CFL condition for this differential equation is satisfied: $c\,dt/h = 0.0025c/0.025 = c/10 < 1/\sqrt{3}$ where $c = \sqrt{\mu(\mathbf{r})/\rho}$ (see below) is between 1 and 2. Absorbing boundary conditions analogous to those in two dimensions (section 4.3.1) are used.

Once again, more than one fixed position (or measurement point) is used to improve the minimisation. They are $r_{ij} = (3, 3 + i, 4 + 2j)/10$, $i = 0, 1, 2, 3, 4$; $j = 0, 1$. The time development of the wave function and its difference from the target reference wave at these points is given in the Figures 4.40 and 4.41. These peaks correspond to each of the ten measurement points. However, in the first plot only five curves are seen corresponding to the original wave arriving at both index $j = 0$ and $j = 1$ of r_{ij} at the same time symmetrically for each index i. Once again each peak position shows when in time the original wave arrive at each measurement point. The highest peak occurs at index $i = 2$ of r_i. Again there is a pairing of the other four peaks, with the higher peak pair at around 0.1 time-units corresponding to arrivals at $i = 1$ and $i = 3$, while the other pair is at around 0.2 time-units for arrivals at $i = 0$ and $i = 4$. No reflections of the feature is seen here. The curves after the peaks fell flat to zero. (In the corresponding plot for two-dimensional Gaussian feature these curves after peaking gradually decreased to, but did not reach, zero, then flattened to a non-zero value after the reflections). There must be reflections, however, but they appear to be overshadowed by interference effects from behind the main wave. The second plot shows the curves are flat for up to 0.3 time-units. Then reflection effects starts to roll in with differences between -1.5×10^{-7} and 2×10^{-7}. The squared difference is up to 4×10^{-14}.

The full wave function over the whole space, projected in a two dimensional

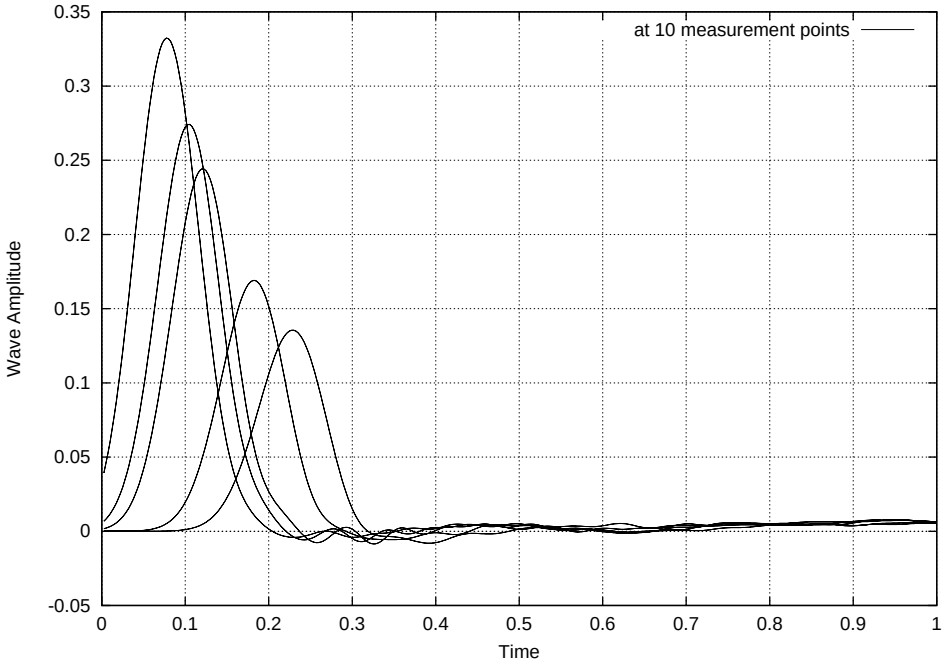

Figure 4.40: The amplitude of time at several measurement points for 3D Gaussian feature

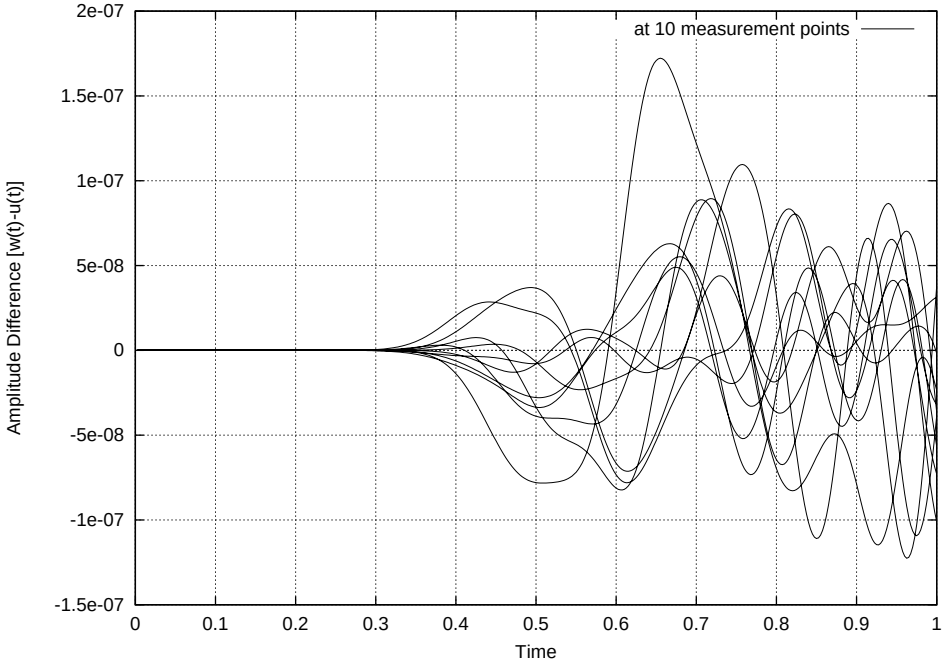

Figure 4.41: The amplitude difference at several measurement points for 3D Gaussian feature

plane at the middle of the third-dimension, follows in Figure 4.42 again with 9 time development frames of wave propagation. These plots is a projection on a two-dimensional xy-plane at the middle of the third-dimension z. The same discussion applies here, as in the case for two-dimensional Gaussian feature. Here the interference behind the main wave is even less pronounced. Once again the waves are not perfect circles in regard of a fewer number of space-steps. This can be increased, but the computer will run into hours of runtime. The amplitude is reduced from unity to 0.008 over the entire time period which made the damping zone less visible except where the wave is absorbed.

Material Property

Here the setup in two-dimensional space above is generalised to three dimensions. λ is represented by a Gaussian feature with the function [mass density $\rho(x, y, z)$ is unity over all space]:

$$\lambda(x, y, z) = \lambda_{env} + (\lambda_{peak} - \lambda_{env}) \exp(-\beta R^2),$$

where $\beta = 320$ is the Gaussian steepness, $R = \sqrt{(x - b_x)^2 + (y - b_y)^2 + (z - b_z)^2}$ is the Gaussian radius, and $\mathbf{b} = (b_x, b_y, b_z) = (0.7, 0.5, 0.5)$ is the position of the feature. λ_{env} is the environment level value and λ_{peak} is the value of the peak.

We use the finite-element method whose coefficients are now calculated from the λ-function above at the corner positions of the elements. The Lagrange polynomials are of fourth order, each operating on rectangular regions (the elements) of a volume $H_x \times H_y \times H_z$. Currently the side-lengths are $\mathbf{H} = (H_x, H_y, H_z) = (6, 6, 6)h$ within each element, which makes them cubic. To accommodate the feature we have 8 cubes, arranged in $2 \times 2 \times 2$, whose common intersecting corners are at position $\mathbf{b}$, in the middle of the applied FEM-region.

The plots for three-dimensional problem follows. The optimised lambda function λ and its difference from its target is seen in Fig.4.43 and 4.44. These plots are a projection on a two-dimensional xy-plane at the middle of the third-dimension z. The first plot shows again a nice spherical Gaussian peak at position $\mathbf{b}$ on the

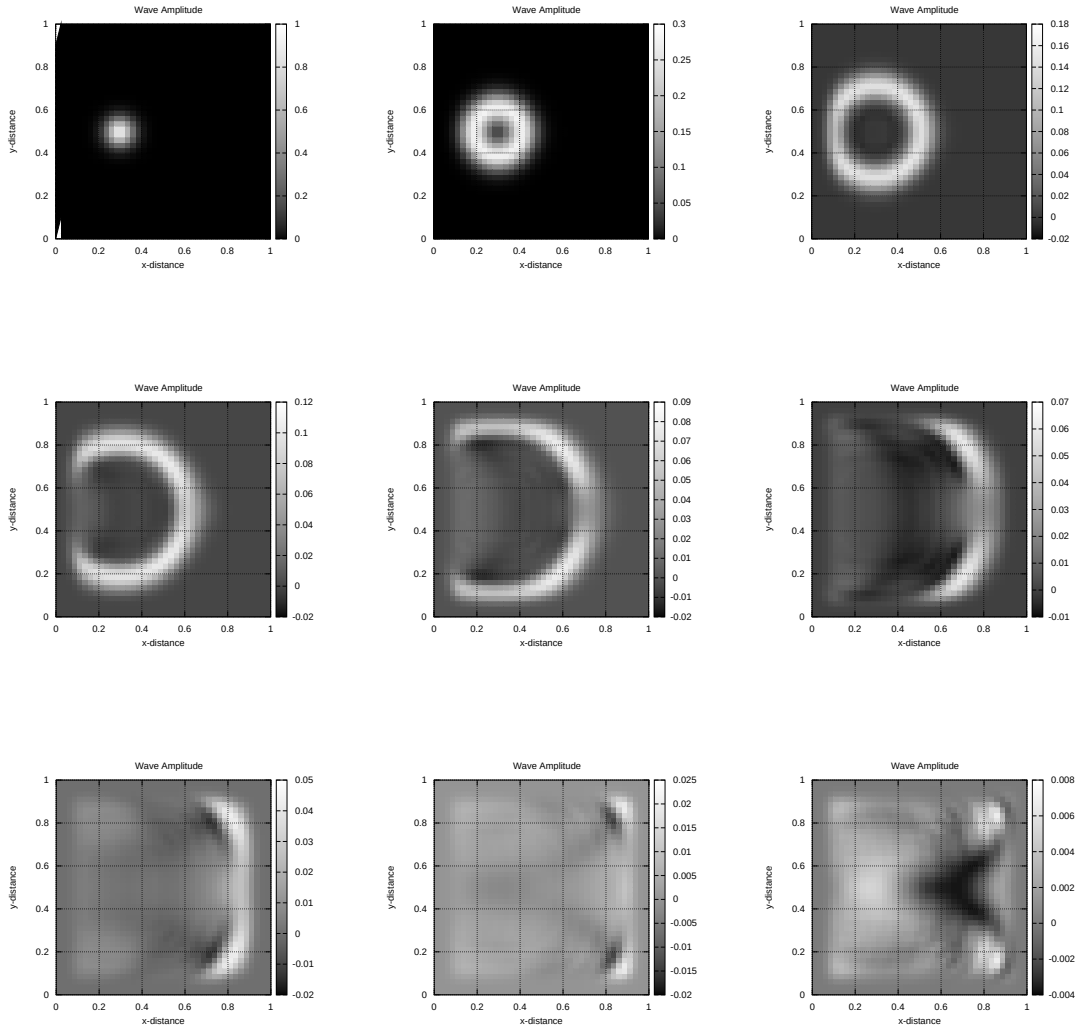

Figure 4.42: The amplitude in 3D space: time frames are from $t = 0$ to $t = 0.8$ in 0.10 seconds increments for 3D Gaussian feature

other side of the space. In the current view of the second plot the differences are more complex. There are four peaks side-by-side (y-direction) and alternating in sign between the negative and positive. A maximum is evident in the middle of this FEM region (y-direction) with one minimum on the near side and two minima on the far side. The differences are in the range of ± 0.012.

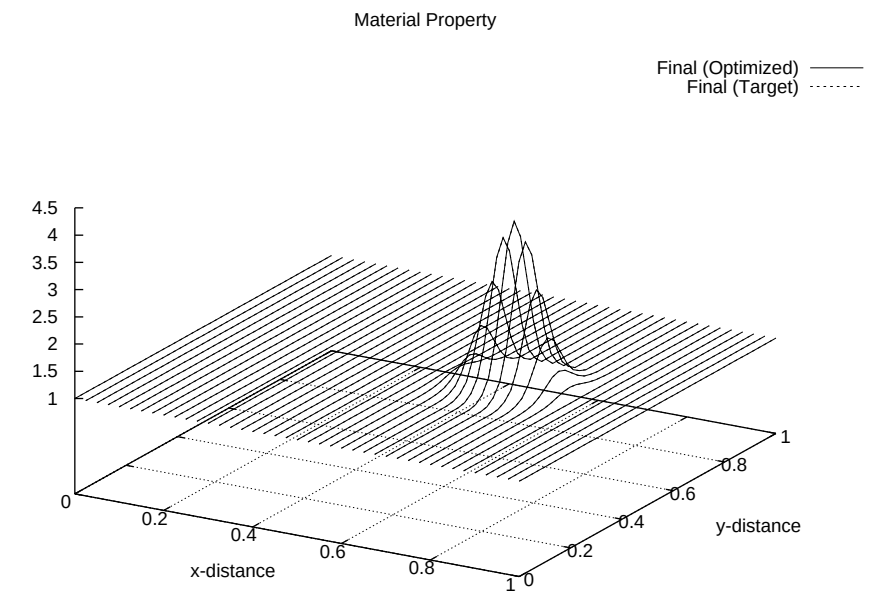

Figure 4.43: The optimised material property function λ for 3D Gaussian feature

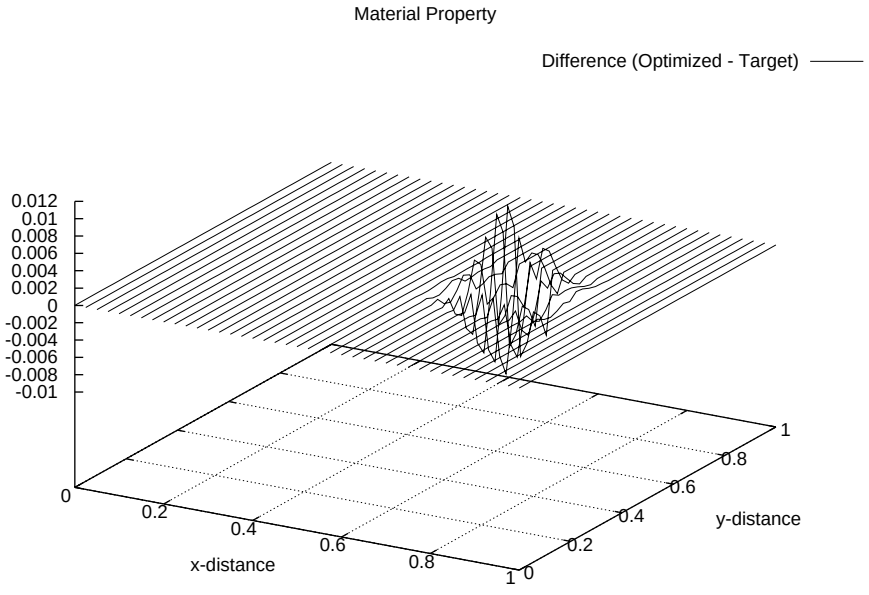

Figure 4.44: The optimised difference to its target function for 3D Gaussian feature

Chapter 5

Conclusions

The property plots in the previous chapter and table 5.1 of relative fraction differences show how the Levenberg-Marquardt minimisation method brings the optimised material property function close to the target model. In the property difference plots, the differences shown are more pronounced at the higher layers, after the transition from the lowest layer has been made. The best results are those for one-dimensional multi-layered interface and two-dimensional feature.

In a comparison between the Fourier series and the finite element methods, the latter method seems more successful. But see an explanation about areas or subregions in section 5.1 below.

The differences of the amplitudes for use in the objective function, as shown in table 5.2, are significantly reduced with the addition of spatial dimensions or interfaces.

5.1 Summary of our results

5.1.1 One dimension

By comparing the material property difference results between the single and multiple interfaces, we see the latter gives a smaller, and thus better, deviation. This is due to the fact that there are more than two layers in the target function which

Table 5.1: Plot ranges of the material property differences of the optimised to the target function (divided by the difference between the lowest λ_o and highest λ_f layer values), values in units of 10^{-2}.

Material property (Method)	Single interface (Fourier series)	Multiple interface (Fourier series)	Feature (Finite elements)
One dimension	[-0.500,0.667]	[-0.100,0.167]	—
Two dimensions	[-5.333,2.000]	[-0.833,1.000]	[-0.133,0.133]
Three dimensions	—	—	[-0.333,0.400]

Table 5.2: Plot ranges of the amplitude differences (the optimised relative to the target reference function) sampled at the measurement points for the objective function, values in units of 10^{-6}.

Material property (Method)	Single interface (Fourier series)	Multiple interface (Fourier series)	Feature (Finite elements)
One dimension	[-600,800]	[-200.0,200.0]	—
Two dimensions	[-25, 10]	[-3.0, 2.5]	[-1.00,4.00]
Three dimensions	—	—	[-0.15,0.20]

approximately comes closer to the initialised optimised function that resembles a piecewise linear function. The minimisation had a closer run to the minimum in the multiple interface setup than that of the single interface. This can also be seen in the amplitude differences.

5.1.2 Two dimensions

Comparisons of the material property differences in two dimensions between the single and multiple interfaces leads to the same conclusion as with one dimensional results as above. The two dimensional results are less satisfactory. The single interface material property result does not even reach our goal of less than or equal to $\pm 10^{-2}$ ($\pm 1\%$) difference (see table 5.1). Since the interface position is at $x = 0.7$, and not $x = 0.5$, the waves reflecting at the interface reaches the measurement points much later and the minimisation did not have much amplitude differences to go by, while the optimised parameter initial point is far from the minimum position in the parameter space. Also, just after the transition from the lower layer, a small trench is visible in the upper layer. For the multi-layer interface results, there are also small differences still visible at the third layer. See section 4.1.2 for an explanation. Also, while in one dimension the real space interval the technique worked on is $[b_o, b_f]$, in two dimensions it is the Cartesian product $[b_o, b_f] \times [0, 1]$ which defines a rectangular strip running under the y-parallel interface where a radially expanding wave impacts on it at various times. The Gaussian feature result is much better than the interfaces due to the feature occupying a smaller area for the FEM region.

5.1.3 Three dimensions

There are no results yet for interface(s) in three dimensions. Thus, for comparison with the fewer dimensions, we have only the Gaussian feature results. Once again the material property deviation results are less accurate than the corresponding two-dimensional feature difference result, which highlights a trend we see across the real space dimensions.

One may ask why it is that the material property difference results become

less and less accurate as one adds one more dimension after another to real space $\mathbb{R}^m$, while the amplitude differences used in the objective function become more accurate. Recall that the more advanced minimisation techniques, such as Levenberg-Marquardt, are designed to work with slightly non-linear mathematical systems (as set out in [11]). Therefore, small changes in one variable or function may lead to large changes in other variables or functions that depend on them. To achieve the same accuracy in the material property difference results ($\pm 1\%$) throughout, we want to redesign the minimisation techniques. That is not yet possible with the current techniques, since they would sometimes cut out after a while with the message that no further progress is possible.

One optimisation design would be to replace the built-in gradient of MERLIN with a more efficient one, so that no plateaus appear in plots of the objective function against minimisation calls (see Figure 4.4 in section 4.1.2).

5.2 Comparison with B.R. Mabuza's work

B.R. Mabuza's work on quantum inverse scattering shows results for square-type and multi-layered potentials through single and coupled channel inversion (chapter 6 in [12]). These results are similar to what we are doing with our material property functions, albeit with a different method.

With his quantum potentials, the deviations are also wavy over the whole space, more pronounced at the transitions from one level to another (and near his boundaries on either side too), but comparable in sizes to what we had found in our results. This means our approach with the Fourier series or finite element method and the minimisation algorithm used, has also been reasonably accurate.

5.3 Comparison of Aims and Results.

In one dimension the target material property (MP) function was defined in section 4.2.1. The corresponding optimised function was designed such that, when its parameters were initialised, it had three piecewise continuous linear functions. A Fourier expansion was used only for the difference between the material property and the linear function. For the single and multiple interface setups the relative MP differences are given in table 5.1. Thus for the one dimensional calculations the resulting accuracy did reach our aim of 1%.

For the two dimensional case our calculations were improved by an increase of the number of measurement points from one to five, ten, or even twenty points at different positions of this space. The single and multiple interface setups were extended parallel to the y-axis in two dimensions. The result of the relative MP difference with the Fourier series method for the single interface is as given in the table. The lowest minimum (see figure 4.25) highlights the trench (see figure 4.24) at the same position. For the multiple interface the relative MP difference is also listed in the table. Figure 4.31 shows oscillating maxima increasing and minima decreasing as one goes towards the opposite space boundary. Thus for the two dimensional interface calculations the resulting accuracy reached our aim of 1% for the multiple interface setup, but not for the single interface setup as evident from this trench.

The two and three dimensional Gaussian features, as optimised with the finite element method, show relative MP differences as given in the table. Figures 4.37 and 4.44 are referred to for more detail. Thus the relative difference is less than for the interface results above. Thus for the two and three dimensional feature calculations the resulting accuracy reached our aim of 1%.

We have yet to develop a more advanced technique for the minimisation of more complex systems and the method as it is, still has to be tested against real data.

5.4 Future Work.

Further research may include more complex physical setups, such as layers not parallel to the y-axis but going off at an oblique angle to it and layers whose interfaces are curvilinear with respect to the y-axis. It may include any combination of the above.

No observed data have been used as input for our methods yet, such as those from survey experiments done on the earth's crust or the sea bottom. We would then be able to fit the model to the given data, and even make alterations to the model for improvements of a fit, but so far the calculations are still only synthetic.

Other possible improvements would be to invert both Lamé parameters, regularise the objective function, place restrictions on the coeffients to make the solution set smaller, and add noise to the synthetic data.

Bibliography

[1] A. Ben-Menahem and S. J. Singh. *Seismic Waves and Sources*. Springer-Verlag New York Incorporated, New York, Heidelberg, Berlin, 1981.

[2] K. Aki and P. G. Richards. *Quantitative Seismology*, volume volume 1 and 2. W.H. Freeman and Company, San Francisco, 1980.

[3] R.J. Atkin and N. Fox. *An introduction to the Theory of Elasticity*. Longman Group Limited, London, 1980.

[4] R. W. Little. *Elasticity*. Prentice-Hall Inc., Englewood, New Jersey, 1973.

[5] H. Leipholz. *Theory of Elasticity*. Noordhoff International Publishing, Leyden, the Netherlands, 1974.

[6] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems*. Springer-Verlag, New York, 1996.

[7] J. Wang and Nicholas Z. Hierarchical bayesian models for inversion problem in heat conduction. *Inverse Problems*, 21:p183–206, 2005. Institute of Physics (IoP), UK.

[8] W. T. Wei, X. X. Heng, and L. R. Xun. Filtering inverse method. *Inverse Problems*, 3:p143–148, 1987. Institute of Physics (IoP), UK.

[9] G. Makrakis and M. I. Taroudakis. *Inverse Problems in Underwater Accoustics*. Springer-Verlag, New York, 2001.

[10] R. Aster and B. Borchers. *Parameter Estimation and Inverse Problems*. Elsevier, London, 2005.

[11] A. Tarantola. *Inverse Problem Theory*. Society of Industrial and Applied Mathematics (SIAM), Philadelphia, 2005. A printed copy.

[12] B. R. Mabuza. *Applied Inverse Scattering*. PhD thesis, University of South Africa, Pretoria, South Africa, November 2005.

[13] E. Polak. *Computational Methods in Optimization: A Unified Approach*, volume 77. Academic Press, New York, London, 1971.

[14] R. Fletcher. *Practical Methods of Optimization*. John Wiley and Sons Limited, New Delhi, 1987.

[15] D.G. Papageorgiou, I.N. Demetropoulos, and I.E. Lagaris. *Merlin 3.0*. University of Ioannina, Greece. A Fortran77 minimization package.

[16] G.A. Evangelakis, J.P. Rizos, I.E. Lagaris, and Demetropoulos. Merlin - a portable system for multdimensional minimization. *Computer Physics Communications*, 46:p401–415, 1987.

[17] A. Tarantola and Valette B. Inverse problems = quest for information. *Geophysics*, 50:p159–170, 1982.

[18] C. Fletcher and M.J.D. Powell. A rapidly convergent descent method for minimization. *Computer Journal*, 6:p163–168, August 1963.

[19] C. Fletcher. A new approach to variable metric algorithms. *Computer Journal*, 13:p317–322, 1970.

[20] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research: National Bureau of Standards*, 49:p409–436, 1952.

[21] J.A. Nedder and R. Meed. A simplex method for function minimization. *Computer Journal*, 7:p308–313, January 1965.

[22] M. I. A. Lourakis. A brief description of the levenberg-marquardt algorithim implemented by `levmar`. Institute of Computer Science, Foundation for Research and Technology, 11 February 2005.

[23] P. Bording. *Seismic Wave Propagation Modeling and Inversion.* Computational Science Educational Project, retrieved on 23 April 2010 at http://www.phy.ornl.gov/csep/CSEP/SW/SW.html.

[24] J. Falk, E. Tessmer, and D. Gajewski. Tube wave modeling by the finite-difference method with varying grid spacing. *Pure and Applied Geophysics*, 148(1/2), March 1996. Birkhauser Verlag, Basel.

[25] G. Cohen and P. Joly. Construction and analysis of fourth order finite-difference schemes for accoustic wave equations in nonhomogeneous media. *SIAM Journal Numerical Analysis*, 33(4):p1266–1302, August 1996. Society of Industrial and Applied Mathematics (SIAM).

[26] A. J. Davies. *The Finite Element Method: A first appraoch.* Clarendon Press, Oxford, 1980.

[27] A.R. Mitchel and R. Wait. *The Finite Element Method in Partial Differential Equations.* John Wiley and Sons, London, New York, Sydney, Toronto, 1977.

[28] R. Wait and A.R. Mitchel. *Finite Element Analysis and Applications.* John Wiley and Sons, Chichester, New York, Brisbane, Toronto, Singapore, 1985.

[29] D. D. Koslof and E. Baysal. Forward modeling by a fourier method. *Geophysics*, 47(10):p1402–1412, October 1982.

[30] A.T. Patera. The spectral element method for fluid dynamics: Laminar flow in a channel expansion. *Journal of Computational Physics*, 54:p468–488, 1984.

[31] D. Komatitsch, J. Vilotte, Rossana Vai, Francisco Sanchez-Sesma, and Jose M. Castillo-Covarrubias. The spectral element method for elastic wave equations – application to 2d and 3d seismic problems. *International Journal for Numerical Methods in Engineering*, 45:p1139–1164, 10 June 1999.

[32] C. Cerjan, M. Reshef, D. Kosloff, and R. Kosloff. A non-reflecting boundary condition for discrete accoustic and elastic wave equations. *Geophysics*, 50(4):p705–708, April 1985.

[33] K. M. T. Kleefsman. *Water Impact Loading on Offshore Structures*. PhD thesis, University of Groningen, November 2005.

[34] O. Vacus. Mathematical analysis of absorbing boundary conditions for the wave equation: the corner problem. *Mathematics of Computation*, 74(249):p177–200, 22 July 2004.

[35] T. Hagstrom. Radiation boundary conditions for the numerical simulation of waves. *Acta Numerica*, 8:p47–106, 1999.

[36] D. Givoli. Review article: Non-reflecting boundary conditions. *Computational Physics*, 94:p1–29, 1991.

[37] M. Braun and J. S. de Villiers. New method for suppressing reflections at the acoustics model boundary. In *Theoretical and Computational Acoustics 2007*, Heraklion, Crete, Greece, 2 to 5 July 2007. University of Crete, Foundation for Research and Thechnology-Hellas.

[38] J. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Computational Physics*, 114:p185–200, 1994.

[39] F. E. Michou and V. K. Koumousis. Absorbing boundary conditions and perfect matched layer: Models for plane soil-structure interaction problems. In *5th GRACM International Congress on Computational Mechanics*, Limassol, 29 June to 1 July 2005. GRACM.

[40] A. Sommerfeld. *Lectures on Theoretical Physics*. Academic Press, New York, 1964.

[41] B. Engquist and A. Majda. Absorbing boundary condition for the numerical simulation of waves. *Journal of Mathematical Computation*, 31(139):p629–651, 1977.

[42] B. Engquist and A. Majda. Radiation boundary conditions for accoustic and elastic waves. *Communication on Pure and Applied Mathematics*, 32:p313–357, 1979.

[43] E. L. Lindman. Free-space boundary conditions for the time-dependent wave equation. *Journal of Computational Physics*, 18:p66–78, 1975.

[44] L. N. Trefethen and L. Halpern. Well-posedness of one-way wave equations and absorbing boundary conditions. *Journal of Mathematical Computation*, 47(176):p421–435, 1986.

[45] L. Halpern and L. N. Trefethen. Wide-angle one-way wave equations. *Accoustic Society of America*, 84:p1397–1404, 1988.

[46] C. J. Randall. Absorbing boundary condition for the elastic wave equation. *Geophysics*, 53(5):p611–624, 1988.

[47] L. Wagatha. Approximation of pseudodifferential operators in absorbing boundary conditions for hyperbolic equations. *Numerical Mathematics Journal*, 42:p51–65, 1983.

[48] A. C. Reynolds. Boundary conditions for the numerical solution of wave propagation problems. *Geophysics*, 43(6):p1099–1110, October 1978.

[49] A. Bayliss and E. Turkel. Radiation boundary condition for wave-like equations. *Communication on Pure and Applied Mathematics*, 33:p707–725, 1980.

[50] A. Bayliss and E. Turkel. Far field boundary conditions for compressible flows. *Journal of Computational Physics*, 48:p182–199, 1982.

[51] A. Bayliss, M. Gunzburger, and E. Turkel. Boundary conditions for the numerical solution of eliptic equations in exterior regions. *SIAM: Applied Mathematics Journal*, 42:p430–451, 1982.

[52] K. Feng. Finite element methods and natural boundary reduction. In *Proceedings. of the International Congress of Mathematicians*, Warsaw, Poland, 1983.

[53] R. L. Higdon. Absorbing boundary conditions for difference approximations of the multi-dimensional wave equation. *Journal of Mathematical Computation*, 47(176):p437–459, 1986.

[54] R. L. Higdon. Numerical absorbing boundary conditions for the wave equation. *Journal of Mathematical Computation*, 49(179):p65–90, 1987.

[55] R. G. Keys. Absorbing boundary conditions for accoustic media. *Geophysics*, 50(6):p892–902, 1985.

[56] J. Lysmer and R. L. Kuhlemeyer. Finite dynamic model for infinite media. *Engineering Mechanical Division*, 95(EM4):p859–877, 1969.

[57] B. Engquyist and L. Halpern. Far-field boundary conditions for computation over long time. *Journal of Applied Numerical Methods*, 4(1):p21–45, 1988.

[58] G. Cohen. *Higher-order Numerical Methods for Transient Wave Equations*. Scientific Computation. Springer-Verlag, Berlin, 2001.

[59] R. Glowinski. *Partial Differential Equations: Modeling and Numerical Simulation (Computational Methods in Applied Sciences*. Springer-Verlag, 2008.

[60] J. Berenger. *Perfectly Matched Layer (PML) for Computational Electromagnetics*. Morgan and Claypool, 2007.