Gradient Calculation of the Traveltime Cost Function without Ray-tracing

Alain Sei and W.W Symes *

June 29, 1994

Abstract

The inverse problem of tomography is an iterative procedure. It requires the computation of the gradient of the traveltime misfit cost function many times. This calculation is customarily done by ray tracing, the path length of the rays being closely related to the gradient. We propose in this work an alternative method to compute the gradient of the traveltime cost function without ray tracing. We use upwind finite difference schemes to compute the traveltime field by solving the eikonal equation. Then by adjoint state techniques we derive a closed-form expression of the gradient of the traveltime cost function. This approach allows an accurate computation of the gradient as well as the freedom to change the norm on the model space.

1 Motivation

Tomography aims at determining the subsurface structure from arrival times picked on prestack seismic data. This method minimizes a cost function which measures the misfit between the computed traveltimes and the traveltime data. The minimization is an iterative procedure and requires the computation of the gradient of the cost function. The computation of the cost function (traveltimes) and gradient with respect to the slowness is customarily done by tracing rays. An argument based on Fermat's principle shows that the traveltime gradient is related to the path length of the rays [3]. But ray tracing is far from routine in complex media [16]. An alternative to traveltime computation consists in solving the eikonal equation directly by finite differences instead of integrating the ray equations. This idea was introduced into the geophysical litterature in [12] and [15] and developped further in [14, 11] and elsewhere. The numerical integration of the eikonal and related equations used in this paper uses upwind schemes [8, 10]. These efficient schemes give accurate solutions of these equations when no turning rays are present. The upwind feature of these schemes corresponds to computing the traveltime in the direction where the rays are out going. This is what a ray tracer does naturally when it computes the rays from the ray equations.

*Department of Computational and Applied Mathematics, Rice University, P.O Box 1892, Houston, Texas 77251-1892,

Since the cost functional in our approach is an algebraic expression in the samples of the solution of the initial and boundary value problem, the adjoint state method is available to compute its gradient. This technique has been used in [4, 2, 13] and elsewhere to construct algorithms for inverse problems for the wave equation. As we shall see it works equally well in the context of traveltime inversion.

This paper concerns the local computation of the traveltime cost functional and its gradient in the following sense: we assume that all rays have everywhere in the computational domain velocity vectors with positive component in one coordinate direction (the positive z direction, to be specific). That is turning rays must not be present. Moreover we assume that the traveltime is given on that part of the boundary (of the computational domain) on which the rays are entering the domain (inflow boundary). That is the part of the boundary on which traveltime data is not given must obey an outflow condition. Both of these restrictions are well-understood necessary conditions for the stability of difference schemes for hyperbolic initial/boundary value problems [8].

Such local traveltime computation schemes may be concatenated to compute the global traveltime field of a point source, without any restriction on turning rays. General techniques to globalize the computation of the traveltime have been developed [15, 14, 11]. Similar globalization is possible for the gradient calculation presented here. The paper is organized as follows. First we introduce the adjoint state and derive the expression of the L^2 and H^1 gradients. The next section is devoted to the numerical method used to compute these quantities. The upwind scheme and its adjoint are described in detail. Then we give an example of such computations and check the accuracy of the gradients versus differential quotient. In the last section we present our conclusions.

2 Problem under consideration

The cost function J measures the misfit between the traveltimes data and the computed traveltimes. Given a slowness field $s(x,z)$ in a domain Ω of R^2, it is defined by:

$$J(s) = \frac{1}{2}\sum_{r\in\mathcal{R}} |\tau(s,r) - t^d(r)|^2 \tag{2.1}$$

where τ are the computed traveltimes, $t^d(r)$ are the traveltime data at the receiver r and $\mathcal{R}$ is the array of receivers. The traveltime τ is linked to the slowness field s by the eikonal equation [1]:

$$\begin{cases} |\nabla\tau|^2 = s^2 & in\ \Omega \\ \\ \tau = \phi & on\ \Gamma_0 \end{cases} \tag{2.2}$$

Here ϕ is the initial profile of the traveltime on the inflow part of the boundary Γ_0. For a point source field we assume that the medium is locally homogeneous around the point source. So we compute analytically the traveltime in a small box around the point source. The traveltimes on the boundary of this initial box become initial profiles for subsequent computations. We assume that the slowness fields we consider are square integrable, which is quite general and includes all applications. To derive the gradient easily we write the cost function J with Dirac masses as follows:

$$J(s) = \frac{1}{2}\int_\Omega \sum_{r\in\mathcal{R}} |\tau(\sigma;x,z) - t^d(x,z)|^2 \,.\, \delta_r(x,z)\, dx\, dz$$

where $\delta_r(x,z)$ is the Dirac mass at the receiver r. A perturbation δs of the slowness field s will give rise to a perturbation $\delta\tau = \tau'(s)\delta s$ of the traveltime τ. We showed in Appendix A that $\delta\tau$ satisfies the following equation:

$$\left\{ \begin{array}{rcll} \nabla\tau.\nabla\delta\tau & = & s.\delta s & in\ \Omega \\ \\ \delta\tau & = & 0 & on\ \Gamma_0 \end{array} \right. \tag{2.3}$$

Therefore the derivative of J with respect to s is given by:

$$J^{'}(s).\delta s \quad = \quad \int_\Omega \delta\tau(x,z) \left\{ \sum_{r\in\mathcal{R}} (\tau(\sigma;x,z) - t^d(x,z)) \,.\, \delta_r(x,z) \right\} \,dx\,dz$$

Let $Res(x,z) = \sum_{r\in\mathcal{R}} (\tau(s;x,z) - t^d(x,z)).\delta_r(x,z)$ be the residuals of the computed traveltimes at the receivers, we have then:

$$J^{'}(s).\delta s \quad = \quad \int_\Omega \delta\tau(x,z)\ Res(s,z)\ dx\ dz \tag{2.4}$$

The definition of the gradient with respect to s, $\nabla J(s) = G(s)$ is linked to the scalar product or the Hilbert space we want to use. The first natural choice is the space of square integrable functions on Ω $L^2(\Omega)$ since we assumed the slownesses to be square integrable. The gradient with respect to the L^2 scalar product is defined by:

$$\int_\Omega \delta s(x,z)\ G(s)(x,z)\ dx\ dz = J^{'}(s).\delta s \tag{2.5}$$

Adjoint state techniques [4, 2, 7] allows us to use integration by parts to make the quantity we desire appear explicitly in the formula. In our case we want to have an expression where δs appears explicitly. So we introduce a new unknown w (the adjoint state) solution of:

$$\left\{ \begin{array}{rcll} -\nabla\left(w\nabla\tau\right) & = & Res & in\ \Omega \\ \\ w & = & 0 & on\ \ \Gamma_1 = \Gamma/\Gamma_0 \end{array} \right. \tag{2.6}$$

The expression of the derivative (2.4) becomes:

$$\begin{array}{rcl} J^{'}(s).\delta s & = & \int_\Omega \delta\tau(x,z)\ Res(x,z)\ dx\ dz \ = \ -\int_\Omega \delta\tau(x,z)\nabla\left(w\nabla\tau\right)\ dx\ dz \\ \\ & = & \int_\Omega (\nabla\tau.\nabla\delta\tau)(x,z)w(x,z)\ dx\ dz \end{array}$$

Since $\delta\tau$ is the solution of (2.3) we have

$$J^{'}(s).\delta s \ = \ \int_\Omega s(x,z)\delta s(x,z)w(x,z)\ dx\ dz \tag{2.7}$$

Now using (2.5) we have:

$$\int_\Omega s(x,z)\delta s(x,z)w(x,z)\,dx\,dz \;=\; \int_\Omega \delta s(x,z)\;G(s)(x,z)\,dx\,dz$$

Since this relation is true for an arbitrary perturbations δs, the L^2 gradient is:

$$G(s)(x,z) \;=\; s(x,z)w(x,z) \qquad (x,z)\in\Omega$$

Another choice for the Hilbert space is suggested by the uncertainty analysis in [5, 6]. They showed that in order to reduce the dependence of the tomographic inverse problem on the discretization the $H^1(\Omega)$ norm on the models (here the slownesses) was a better choice than the L^2 norm. In that case the gradient verifies:

$$\int_\Omega \delta s(x,z)G(s)(x,z)+\nabla\delta s(x,z)\nabla G(s)(x,z)\,dxdz = J^{'}(s).\delta s \tag{2.8}$$

With (2.7) we can write:

$$J^{'}(s).\delta s = \int_\Omega \delta s(x,z)s(x,z)w(x,z)\,dx\,dz$$

$$= \int_\Omega \delta s(x,z)G(s)(x,z)\,dx\,dz + \int_\Omega \nabla\delta s(x,z)\;\nabla G(s)(x,z)\,dx\,dz$$

$$= \int_\Omega \delta s(x,z)[G(s)(x,z)-\Delta G(s)(x,z)]\,dx\,dz + \int_\Gamma \frac{\partial G}{\partial n}\delta s(x,z)\,d\sigma$$

Therefore the H^1 gradient is the solution of the following elliptic problem:

$$\left\{\begin{array}{rcll} G-\Delta G & = & s\;w & \mathit{in}\ \Omega \\ \dfrac{\partial G}{\partial n} & = & 0 & \mathit{on}\ \Gamma \end{array}\right. \tag{2.9}$$

3 The Discrete Problem

For the computation of the traveltime τ, we have chosen the second order essentially non-oscillatory (ENO) scheme described explicitly in [10]. It is quite efficient and very accurate: for the examples used here, errors in τ are a small fraction of 1%.

The key to the computation of the L^2 or H^1 gradient is the adjoint state w. It is determined by equation (2.3) for the perturbation of the traveltime $\delta\tau$ (which we call the direct equation). In the discrete case, the choice of a numerical scheme for the integration of (2.3) determines the numerical scheme for the integration of the adjoint equation. It is *not* just any discretization of the adjoint equation.

The equation verified by the traveltime perturbation $\delta\tau$ can be written as follows:

$$\left\{\begin{array}{rcll} \tau_z\,\delta\tau_z\;+\;\tau_x\,\delta\tau_x & = & s.\delta s & \mathit{in}\ \ \Omega \\ \delta\tau & = & 0 & \mathit{on}\ \ \Gamma_0 \end{array}\right.$$

For our study we choose the domain to be a rectangle $\Omega =]x_{min}, x_{max}[\times]z_{min}, z_{max}[$ and the boundary Γ_0 to be $z = z_{min}$. We also assume that the slowness field we consider do not produce turning rays. Therefore the derivative of the traveltime τ with respect to z does not vanish ($\tau_z \neq 0$). So we can write the preceeding equation as an evolution equation in z. For convenience we set $u = \delta\tau$ we have:

$$
(3.1) \qquad \begin{cases} u_z + \dfrac{\tau_x}{\tau_z} u_x = \dfrac{s}{\tau_z} . \delta s & in\ \Omega \\ u = 0 & on\ \Gamma_0 \end{cases}
$$

The numerical method we choose to integrate this equation is a first order upwind scheme [8] given by :

$$
(3.2) \qquad \begin{cases} \dfrac{u_j^{n+1} - u_j^n}{\Delta z} - (a^+)_j^n \dfrac{u_{j-1}^n - u_j^n}{\Delta x} + (a^-)_j^n \dfrac{u_{j+1}^n - u_j^n}{\Delta x} = f_j^n & \\ u_j^1 = 0 & j = 1..J \end{cases}
$$

where $a = \dfrac{\tau_x}{\tau_z}$, $a^+ = max(a, 0)$, $a^- = min(a, 0)$ and $f = \dfrac{s}{\tau_z}.\delta s$.

Of course this not the pertubational equation of the second order ENO scheme used to compute τ. Thus the computed derivative $\delta\tau$ and ultimately the gradient derived from it, will disagree with the true gradient of the discrete cost function by a truncation error. As we shall see this error can be made arbitrarily small by refining the grid. More precise gradient schemes can be developed at the cost of more involved formulae. The reason to choose the first order scheme lies in the simplicity of the discrete adjoint equation.

This scheme works if there is "outflow" at the boundaries. The necessary outflow condition on τ translates here into :

$$
\begin{cases} (a^+)_1^n = 0 & (a^+)_2^n = 0 & n = 1..N \\ (a^-)_J^n = 0 & (a^-)_{J-1}^n = 0 & n = 1..N \end{cases}
$$

The Von Neumann stability criterion shows that Δx and Δz can not be chosen independently of one another, but must satisfy:

$$
\left| \frac{a \Delta z}{\Delta x} \right| \leq 1
$$

Appendix B shows that the discrete adjoint state w is the solution of :

$$
(3.3) \qquad \begin{cases} \dfrac{w_j^{n-1} - w_j^n}{\Delta z} - \dfrac{(a^+)_{j+1}^n w_{j+1}^n - (a^+)_j^n w_j^n}{\Delta x} + \dfrac{(a^-)_{j-1}^n w_{j-1}^n - (a^-)_j^n w_j^n}{\Delta x} = Res_j^n \\ w_1^n = w_J^n = 0 \qquad n = 2..N \\ w_j^{N-1} = 0 \qquad j = 1..J \end{cases}
$$

We are now ready to compute the discrete gradients. Equation (2.4) is approximated by the following expression:

$$
(J'(s).\delta s)_h = \sum_{n=1}^{N-1} \sum_{j=2}^{J-1} u_j^n \ Res_j^n \ \Delta x \ \Delta z
$$

Therefore using (B.1) (see Appendix B) we find:

$$(J'(s).\delta s)_h \;=\; \sum_{n=1}^{N-1}\sum_{j=2}^{J-1} w_j^n \,\frac{s_j^n}{(\tau_z)_j^n}(\delta s)_j^n \;\Delta x\;\Delta z$$

whence the L^2 gradient (G_j^n):

$$G_j^n \;=\; w_j^n \,\frac{s_j^n}{(\tau_z)_j^n} \qquad j=2,J-1 \quad n=1,N-1 \tag{3.4}$$

For the H^1 case we must choose an approximation to ∇u. We decide to use a first order approximation. Therefore we replace the gradient by the following approximation $(\nabla_h u)$:

$$(\nabla_h u)_j^n \;=\; \left(\frac{u_{j+1}^n - u_j^n}{\Delta x}, \frac{u_j^{n+1} - u_j^n}{\Delta z}\right)$$

The discrete H^1 scalar product is therefore:

$$\sum_{n=2}^{N-1}\sum_{j=2}^{J-1} u_j^n v_j^n \;\Delta x\;\Delta z + \sum_{n=2}^{N-1}\sum_{j=2}^{J-1}(\nabla_h u)_j^n.(\nabla_h v)_j^n \;\Delta x\;\Delta z$$

Now we can find the H^1 adjoint. Using (B.1) again we can write:

$$\sum_{n=2}^{N-1}\sum_{j=2}^{J-1} u_j^n Res_j^n \;\Delta x\;\Delta z = \sum_{n=2}^{N-1}\sum_{j=2}^{J-1} w_j^n \,\frac{s_j^n}{(\tau_z)_j^n}\delta s_j^n \;\Delta x\;\Delta z$$

$$= \sum_{n=2}^{N-1}\sum_{j=2}^{J-1} G_j^n \delta s_j^n + \frac{G_{j+1}^n - G_j^n}{\Delta x}\,\frac{\delta s_{j+1}^n - \delta s_j^n}{\Delta x} + \frac{G_j^{n+1} - G_j^n}{\Delta z}\,\frac{\delta s_j^{n+1} - \delta s_j^n}{\Delta z} \;\Delta x\;\Delta z$$

By discrete integration by parts (see Appendix C) we find that the discrete H^1 gradient (G_j^n) is the solution of the following problem:

$$\left\{\begin{array}{llll} G_j^n - \dfrac{G_{j+1}^n - 2G_j^n + G_{j-1}^n}{\Delta x^2} - \dfrac{G_j^{n+1} - 2G_j^n + G_j^{n-1}}{\Delta z^2} = w_j^n \,\dfrac{s_j^n}{(\tau_z)_j^n} & & & \\ \dfrac{G_j^N - G_j^{N-1}}{\Delta z} = 0 & j=1..J-1 & \dfrac{G_j^2 - G_j^1}{\Delta z} = 0 & j=1..J-1 \\ \dfrac{G_J^n - G_{J-1}^n}{\Delta x} = 0 & n=1..N-1 & \dfrac{G_2^n - G_1^n}{\Delta x} = 0 & n=1..N-1 \end{array}\right. \tag{3.5}$$

4 Numerical Results

We consider the domain $\Omega =]-1,1[\times]1,3[$. The source of the wave is located at the origin $(0,0)$. The receivers are located at the depth $z=3$ symetrically with respect to $x=0$ (see figure 1).

4.1 Gradient Calculations

The traveltime data we used have been generated with the slowness field $s + ds$, where s is the constant slowness field equal to 1 everywhere and ds is the perturbation displayed in figure 2. The traveltime data displayed in figure 3 show that the slowness field created a caustic. It is apparent in the non smoothness of the traveltime field. This shows that our method works in the presence of caustics. The figures 4 and 5 are the L^2 and H^1 gradients of the cost function computed at the slowness field s. Figure 4 and 5 show that the two gradients have very different smoothness features. This is expected because the H^1 gradient is the solution of an elliptic problem whose source is the L^2 gradient.
The next iterate of the slowness chosen by the optimization algorithm will be in the direction opposite to the gradient. In this example we see that it will increase the slowness from the constant value $s = 1$. That is consistent with the data which have been generated with the slowness $1 + \delta s$ (δs is displayed in figure 2).

4.2 Verification of the gradient by finite differences

Using the definition of the derivative, the gradient of the cost function with respect to the L^2 or H^1 norms can be checked by finite differences. In fact for an arbitrary perturbation in the slowness field δs and a number h we can write using the centered finite difference approximation:

$$\frac{J(s+h.\delta s) - J(s-h.\delta s)}{2h} = J'(s).\delta s + O(h^2) \qquad when \ \ h \longrightarrow 0$$

Since we can replace $J'(s).\delta s$ by $(G(s), \delta s)_H$ where H is either L^2 or H^1 we can test the accuracy of the gradient calculations simply by plotting the curve:

$$h \longmapsto \frac{J(s+h.\delta s) - J(s-h.\delta s)}{2h} - (G(s), \delta s)_H$$

for arbitrary perturbations δs. If it is a parabolic function of the parameter h then we will validate our calculations. As pointed out above the $h = 0$ limit in both cases will generally be non zero as we have not computed the actual discrete gradient, but only an approximation differing by a truncation error.
This is confirmed in figure 6 were we have computed the gradients with three different grids. The grid used for the $*$ curve is four times finer than the grid used for the $+$ curve which is four times finer than the grid used for the $\times$ curve. We see as we refine the grid that the difference between the actual discrete gradient and its approximation tends to zero. The slowness perturbation used in the computation of the finite difference quotients is given in figure 2.

We also verified the convergence on the case of a velocity linearly increasing with depth. We used the perturbation shown in figure 2 to compute the differential quotients. The results displayed in figure 7 show the convergence of the method in that case as well.

5 Conclusions

In this paper we derived a method to compute the gradient of the traveltime cost function without ray tracing. Our approach is based on adjoint state techniques. It consists in computing the adjoint state by retro-propagation of the residuals traveltimes. Then according to the norm chosen on the model space (here the slownesses) we derive a closed form formula for the gradient. This flexibility in the choice of the norm is essential in the numerical treatment of the inverse problem. The numerical method we used is based on upwind schemes. These efficient schemes are limited by their upwind feature. In our case it meant that no turning rays were present in the medium. This, however, does not mean that caustics are absent. In fact our method deals gracefully with caustics as is demonstrated in the preceeding example. Moreover the local calculation explained here may be embedded in a globalization scheme to remove the restrictions of turning rays and outflow boundaries.

Acknowlegement

This work was partially supported by the Texas Geophysical Parallel Computation Project, the Schlumberger Foundation, and The Rice Inversion Project. TRIP Sponsors for 1994 are Advance Geophysical, Amoco Production Co., Conoco Inc., Cray Research Inc., Exxon Production Research Co., Interactive Network Technologies, Mobil Research and Development Corp., and Texaco Inc.

References

[1] K. Aki and P.G Richards. *Quantitative Seismology.* Freeman and Co, 1980.

[2] A. Bamberger, G. Chavent, Ch. Hemon, and P. Lailly. Inversion of normal incidence seiemogramms. *Geophysics*, 47:757–770, 1979.

[3] T. Bishop, K. Bube, R. Cutler, R. Langan, P. Love, J. Resnick, R. Shuey, D. Spindler, and H. Wyld. Tomographic determination of velocity profile and depth in laterally varying media. *Geophysics*, 50:903–923, 1985.

[4] G. Chavent. Identification of functional parameters in partial differential equations. *Identification of parameter distributed systems: R.E. Goodson, and Polis, Ed., New York, ASME*, 1974.

[5] F. Delprat-Jannaud and P. Lailly. What information on the earth model do reflection travel times provide? *Journal of Geophysical Resarch*, 97:19827–19844, 1992.

[6] F. Delprat-Jannaud and P. Lailly. Ill-posed and well-posed formulations of the reflection travel time tomography problem. *Journal of Geophysical Resarch*, 98:6589–6605, 1993.

[7] P. Kolb, F. Collino, and P. Lailly. Prestack inversion of a 1d medium. *Proceedings of the IEEE*, 74:498–508, 1986.

[8] R.J. Leveque. *Numerical Methods for Conservation Laws.* Birkhauser Verlag, Basel, 1992.

[9] R.K. Luneburg. *Mathematical Theory of Optics.* University of California Press, 1964.

[10] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithm based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.

[11] P. Podvin and I. Lecomte. Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. *Geophys. J. Int.*, 105:271–284, 1991.

[12] M. Reshef and D. Kosloff. Migration of common shot gathers. *Geophysics*, 51:324–331, 1986.

[13] A. Tarantola. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49:1259–1266, 1984.

[14] J. VanTrier and William W. Symes. Upwind finite-difference calculation of traveltimes. *Geophysics*, 56:812–821, 1991.

[15] J. Vidale. Finite-difference calculation of travel times. *Bulletin of the Seismological Society of America*, 78:2062–2076, 1988.

[16] J. Virieux and V. Farra. Ray tracing in 3-d complex isotropic media: An analysis of the problem. *Geophysics*, 56:2057–2069, 1991.

A Derivation of the perturbed equation

We consider a slowness perturbation δs. The travel time $\tau(s+\delta s)$ associated to the perturbation satisfies :

$$\begin{cases} \nabla\tau(s+\delta s)^2 = (s+\delta s)^2 & in\ \Omega \\ \tau(s+\delta s) = \phi & on\ \Gamma_0 \end{cases}$$

We want to find the equation satisfied by $\delta\tau = \tau^{'}(s).\delta s$. Since we have

$$\begin{aligned} (\nabla\tau(s+\delta s))^2 &= (\nabla(\tau(s)+\tau^{'}(s).\delta s + o(\delta s^2)))^2 \\ &= (\nabla\tau(s))^2 + 2.\nabla\tau(s).\tau^{'}(s).\delta s + o(\delta s^2)) \end{aligned}$$

we can write

$$2.\nabla\tau(s).\tau^{'}(s).\delta s + o(\delta s^2)) = (\nabla\tau(s+\delta s))^2 - (\nabla\tau(s))^2 = 2s.\delta s + o(\delta s^2)$$

Dropping the term of order higher or equal to two (because we are looking for the first derivative) we find that $\delta\tau$ is the solution of the following problem :

$$\text{(A.1)} \qquad \begin{cases} \nabla\tau.\nabla\delta\tau(x,z) = s.\delta s(x,z) & (x,z)\in\Omega \\ \delta\tau(x,z) = 0 & (x,z)\in\Gamma_0 \end{cases}$$

Remark

Using the method of characteristics we can find from equation (A.1) the usual formula for $\delta\tau$ given by:

$$\delta\tau = \int_0^t \delta s(x(\sigma),z(\sigma))d\sigma$$

where $(x(\sigma),z(\sigma))$ are the rays and σ the arclength.
Since the slowness s is never equal to zero, we can rewrite equation (A.1) as follows:

$$\text{(A.2)} \qquad \begin{cases} \delta\tau_x\frac{\tau_x}{s} + \delta\tau_z\frac{\tau_z}{s} = \delta s & (x,z)\in\Omega \\ \delta\tau(x,z) = 0 & (x,z)\in\Gamma_0 \end{cases}$$

The characterics of equation (A.2) are the curves $(a(t),b(t),\delta\tau(t))$ such that

$$\frac{da}{dt} = \frac{\tau_x}{s} \qquad \frac{db}{dt} = \frac{\tau_z}{s} \qquad \frac{d\delta\tau}{dt} = \delta s$$

Therefore $\delta\tau$ is given by :

$$\delta\tau = \int_0^t \delta s(a(\sigma),b(\sigma))\ d\sigma$$

Furthermore we have :

$$\begin{cases} \dfrac{d}{dt}\left(s\dfrac{da}{dt}\right) = \dfrac{d\tau_x}{dt} = \dfrac{1}{s}\left(\dfrac{1}{2}\dfrac{\partial}{\partial x}(\tau_x^2+\tau_z^2)\right) = s_x \\ \dfrac{d}{dt}\left(s\dfrac{db}{dt}\right) = \dfrac{d\tau_z}{dt} = \dfrac{1}{s}\left(\dfrac{1}{2}\dfrac{\partial}{\partial z}(\tau_x^2+\tau_z^2)\right) = s_z \\ \left(\dfrac{da}{dt}\right)^2 + \left(\dfrac{db}{dt}\right)^2 = \left(\dfrac{\tau_x}{s}\right)^2 + \left(\dfrac{\tau_z}{s}\right)^2 = 1 \end{cases}$$

The last equation means that the parameter t is in fact the arclength and the two first equations are the differential equations satisfied by the rays [9]. Therefore once again we find that the perturbation of the travel time is the integral along the rays (the characteristics) of the pertubation in slowness.

B Adjoint Upwind Scheme

We are looking for the adjoint equation of equation (3.2) for the discrete L^2 scalar product. We use the notation $(.,.)_h$ for that scalar product. Let us note P^* the adjoint operator of P where P is the defined by the discrete equation (3.2). We have

$$(P^*w,u)_h = (Pu,w)_h$$

$$= \sum_{j=2}^{J-1}\sum_{n=1}^{N-1} w_j^n \left(\frac{u_j^{n+1}-u_j^n}{\Delta z} - (a^+)_j^n \frac{u_{j-1}^n - u_j^n}{\Delta x} + (a^-)_j^n \frac{u_{j+1}^n - u_j^n}{\Delta x}\right)\Delta x \Delta z$$

Let us treat the first integral

$$\begin{aligned} I_1 &= \sum_{j=2}^{J-1}\sum_{n=1}^{N-1} w_j^n \frac{u_j^{n+1}-u_j^n}{\Delta z}\Delta x\Delta z \\ &= \sum_{j=2}^{J-1}\frac{1}{\Delta z}\left(\sum_{n=1}^{N-1} w_j^n u_j^{n+1} - \sum_{n=1}^{N-1} w_j^n u_j^n\right)\Delta x\Delta z \\ &= \sum_{j=2}^{J-1}\frac{1}{\Delta z}\left(\sum_{n=2}^{N} w_j^{n-1} u_j^{n} - \sum_{n=1}^{N-1} w_j^n u_j^n\right)\Delta x\Delta z \\ &= \sum_{j=2}^{J-1}\sum_{n=1}^{N-1} \frac{w_j^{n-1}-w_j^n}{\Delta z} u_j^n \Delta x\Delta z - + \sum_{j=2}^{J-1} w_j^{N-1}u_j^N \Delta x \qquad \textit{since}\ \ u_j^1 = 0 \end{aligned}$$

The second integral can be written as

$$\begin{aligned} I_2 &= -\sum_{j=2}^{J-1}\sum_{n=1}^{N-1} w_j^n (a^+)_j^n \frac{u_{j-1}^n - u_j^n}{\Delta x}\Delta x\Delta z \\ &= -\sum_{n=1}^{N-1}\frac{1}{\Delta x}\left(\sum_{j=2}^{J-1}(a^+)_j^n w_j^n u_{j-1}^n - \sum_{j=1}^{J-1}(a^+)_j^n w_j^n u_j^n\right)\Delta x\Delta z \\ &= -\sum_{n=1}^{N-1}\frac{1}{\Delta x}\left(\sum_{j=1}^{J-2}(a^+)_{j+1}^n w_{j+1}^n u_j^n - \sum_{j=2}^{J-1}(a^+)_j^n w_j^n u_j^n\right)\Delta x\Delta z \end{aligned}$$

$$= -\sum_{j=2}^{J-1}\sum_{n=1}^{N-1}\frac{(a^+)_{j+1}^n w_{j+1}^n - (a^+)_j^n w_j^n}{\Delta x} u_j^n \Delta x\Delta z + \sum_{n=1}^{N-1}(a^+)_J^n w_J^n u_{J-1}^n \Delta z \qquad since \ \ (a^+)_2^n = 0$$

Let us treat the last integral

$$\begin{aligned} I_3 &= \sum_{j=2}^{J-1}\sum_{n=1}^{N-1} w_j^n (a^-)_j^n \frac{u_{j+1}^n - u_j^n}{\Delta x}\Delta x\Delta z \\ &= \sum_{n=1}^{N-1}\frac{1}{\Delta x}\left(\sum_{j=2}^{J-1}(a^-)_j^n w_j^n u_{j+1}^n - \sum_{j=2}^{J-1}(a^+)_j^n w_j^n u_j^n\right)\Delta x\Delta z \\ &= \sum_{n=1}^{N-1}\frac{1}{\Delta x}\left(\sum_{j=3}^{J}(a^-)_{j-1}^n w_{j-1}^n u_j^n - \sum_{j=2}^{J-1}(a^-)_j^n w_j^n u_j^n\right)\Delta x\Delta z \\ &= \sum_{j=2}^{J-1}\sum_{n=1}^{N-1}\frac{(a^-)_{j-1}^n w_{j-1}^n - (a^-)_j^n w_j^n}{\Delta x} u_j^n \Delta x\Delta z - \sum_{n=1}^{N-1}(a^-)_1^n w_1^n u_2^n \Delta z \qquad since \ \ (a^-)_{J-1}^n = 0 \end{aligned}$$

Finally we can write

$$\begin{aligned} &\sum_{j=2}^{J-1}\sum_{n=1}^{N-1}\left(\frac{w_j^{n-1} - w_j^n}{\Delta z} - \frac{(a^+)_{j+1}^n w_{j+1}^n - (a^+)_j^n w_j^n}{\Delta x} + \frac{(a^-)_{j-1}^n w_{j-1}^n - (a^-)_j^n w_j^n}{\Delta x}\right) u_j^n \Delta x\Delta z \\ &= \sum_{j=2}^{J-1}\sum_{n=1}^{N-1} w_j^n\left(\frac{u_j^{n+1} - u_j^n}{\Delta z} - (a^+)_j^n\frac{u_{j-1}^n - u_j^n}{\Delta x} + (a^-)_j^n\frac{u_{j+1}^n - u_j^n}{\Delta x}\right)\Delta x\Delta z \\ &+ \sum_{n=1}^{N-1}(a^-)_1^n w_1^n u_2^n \Delta z - \sum_{j=1}^{J-1} w_j^{N-1} u_j^N \Delta x - \sum_{n=1}^{N-1}(a^+)_J^n w_J^n u_{J-1}^n \Delta z \end{aligned}$$

So if we choose w such that :

$$\begin{cases} w_1^n = w_J^n = 0 & n = 2..N-1 \\ w_j^{N-1} = 0 & j = 2..J-1 \end{cases}$$

we find that the adjoint scheme of (3.2) is given by (3.3). Using the right handside of equation (3.2) and equation (3.3) we can write

$$\sum_{n=1}^{N-1}\sum_{j=2}^{J-1} u_j^n \; Res_j^n \; \Delta x \; \Delta z \; = \; \sum_{n=1}^{N-1}\sum_{j=2}^{J-1} w_j^n \; \frac{s_j^n}{(\tau_z)_j^n}(\delta s)_j^n \; \Delta x \; \Delta z \tag{B.1}$$

C Discrete H^1 gradient

We have

$$\begin{aligned} &\sum_{n=2}^{N-1}\sum_{j=2}^{J-1} w_j^n \; \frac{s_j^n}{(\tau_z)_j^n}\delta s_j^n \; \Delta x \; \Delta z \\ &= \sum_{n=2}^{N-1}\sum_{j=2}^{J-1} G_j^n \delta s_j^n + \frac{G_{j+1}^n - G_j^n}{\Delta x}\frac{\delta s_{j+1}^n - \delta s_j^n}{\Delta x} + \frac{G_j^{n+1} - G_j^n}{\Delta z}\frac{\delta s_j^{n+1} - \delta s_j^n}{\Delta z} \; \Delta x \; \Delta z \end{aligned} \tag{C.1}$$

To derive the equation satisfied by G_j^n we have to factor out δs_j^n. The first term is not a problem since δs_j^n is already factored out. For the second and third term we have to integrate by parts. We show the calculation on second term, the third term is treated exactly in the same fashion.

$$\begin{aligned} &\sum_{n=2}^{N-1}\sum_{j=2}^{J-1} \frac{G_{j+1}^n - G_j^n}{\Delta x}\frac{\delta s_{j+1}^n - \delta s_j^n}{\Delta x} \; \Delta x \; \Delta z \\ &= \sum_{n=2}^{N-1} \frac{1}{\Delta x}\left(\sum_{j=2}^{J-1} \delta s_{j+1}^n \frac{G_{j+1}^n - G_j^n}{\Delta x} - \delta s_j^n \frac{G_{j+1}^n - G_j^n}{\Delta x}\right) \; \Delta x \; \Delta z \\ &= \sum_{n=2}^{N-1} \frac{1}{\Delta x}\left(\sum_{j=3}^{J} \delta s_j^n \frac{G_j^n - G_{j-1}^n}{\Delta x} - \sum_{j=2}^{J-1} \delta s_j^n \frac{G_{j+1}^n - G_j^n}{\Delta x}\right) \; \Delta x \; \Delta z \\ &= -\sum_{n=2}^{N-1}\sum_{j=2}^{J-1} \frac{G_{j+1}^n - 2G_j^n + G_{j-1}^n}{\Delta x^2}\delta s_j^n \; \Delta x \; \Delta z - \sum_{n=2}^{N-1} \delta s_2^n \frac{G_2^n - G_1^n}{\Delta x} \; \Delta z + \sum_{n=2}^{N-1} \delta s_J^n \frac{G_J^n - G_{J-1}^n}{\Delta x} \; \Delta z \end{aligned}$$

Identically we have:

$$\sum_{n=2}^{N-1}\sum_{j=2}^{J-1}\frac{G_j^{n+1}-G_j^n}{\Delta x}\frac{\delta s_j^{n+1}-\delta s_j^n}{\Delta x}\ \Delta x\ \Delta z$$

$$= -\sum_{n=2}^{N-1}\sum_{j=2}^{J-1}\frac{G_j^{n+1}-2G_j^n+G_j^{n-1}}{\Delta x^2}\delta s_j^n\ \Delta x\ \Delta z - \sum_{j=2}^{J-1}\delta s_j^2\frac{G_j^2-G_j^1}{\Delta x}\ \Delta z + \sum_{j=2}^{J-1}\delta s_j^N\ \frac{G_j^N-G_j^{N-1}}{\Delta x}\ \Delta z$$

Therefore according to (C.1) G_j^n is solution of (3.5).

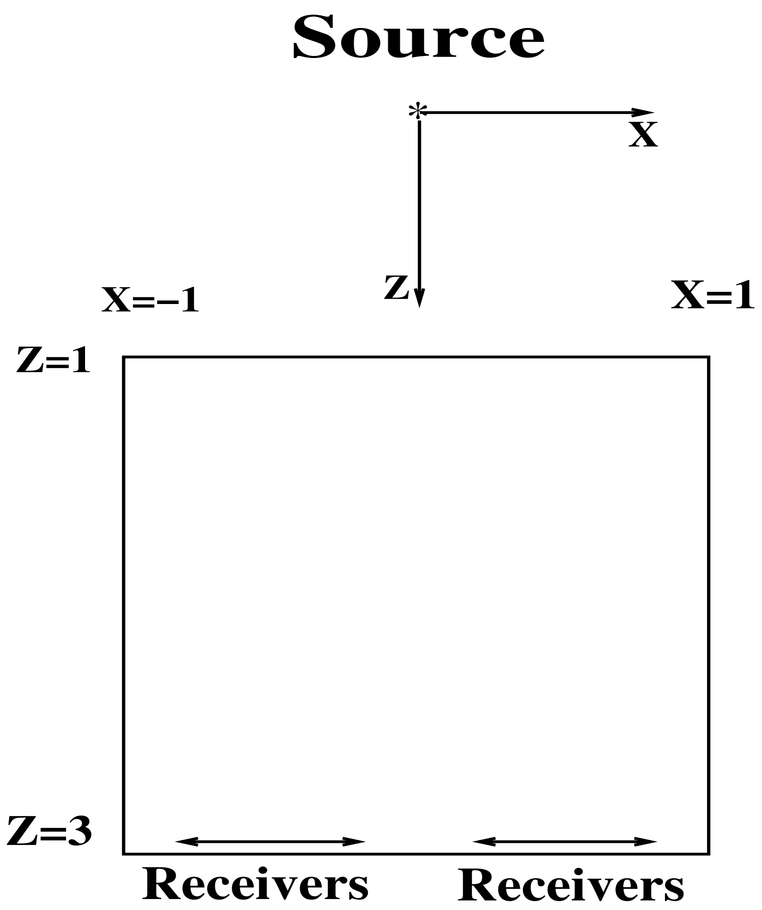

Figure 4.1: The Medium and the Source Receiver Configuration

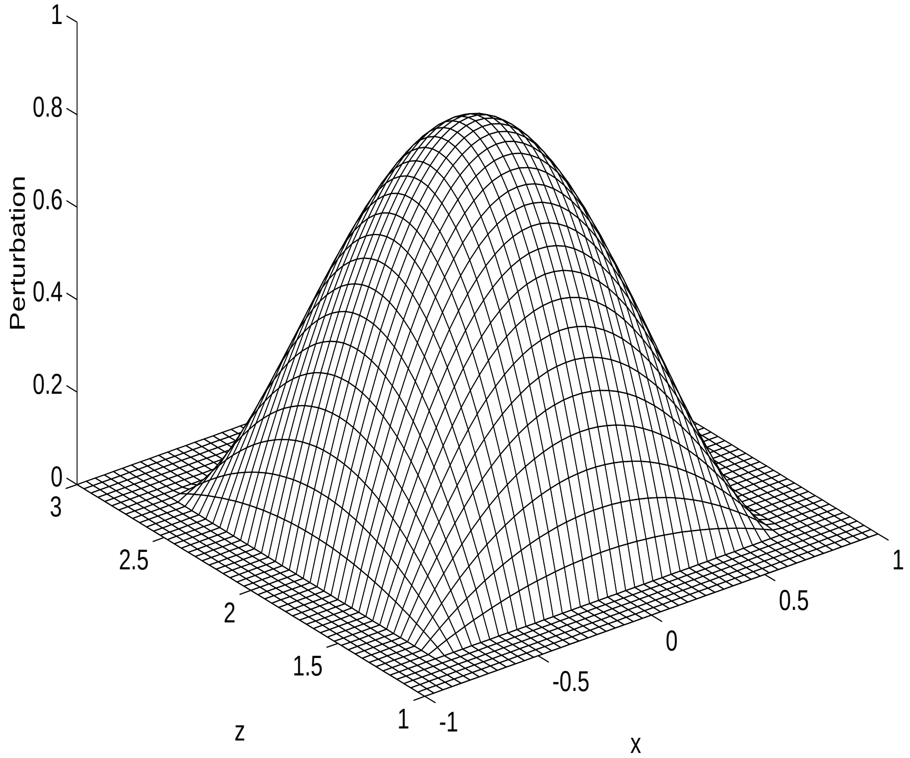

Figure 4.2: The slowness perturbation

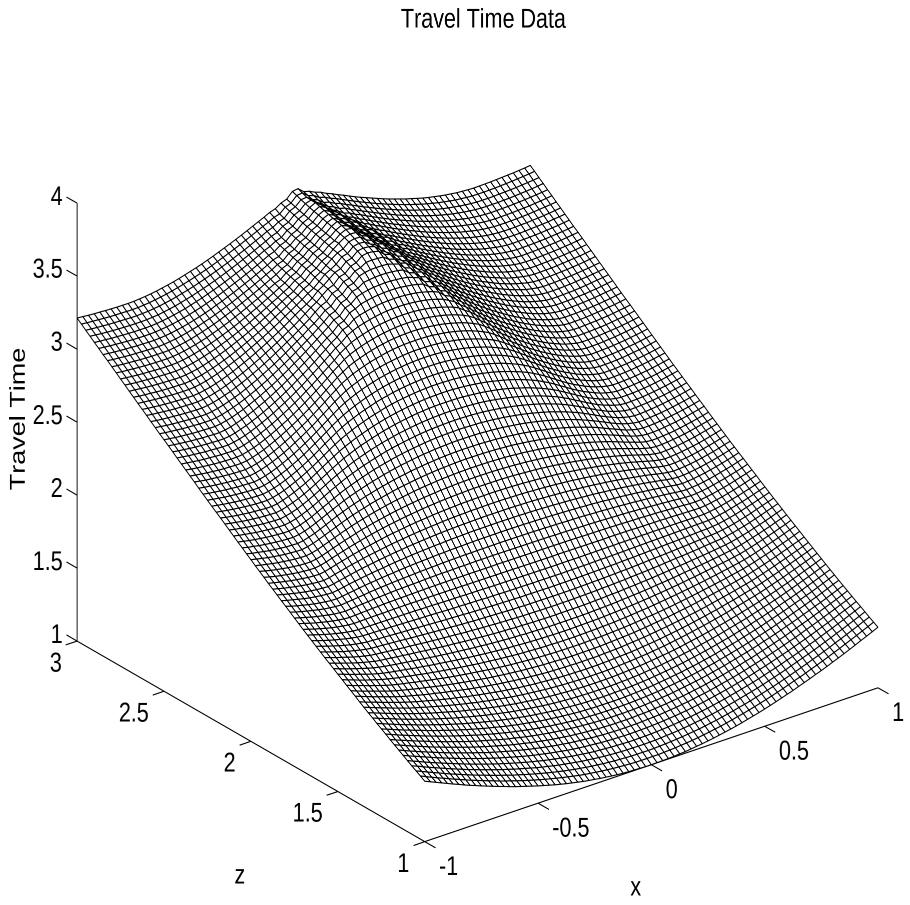

Figure 4.3: Traveltime data with a caustic

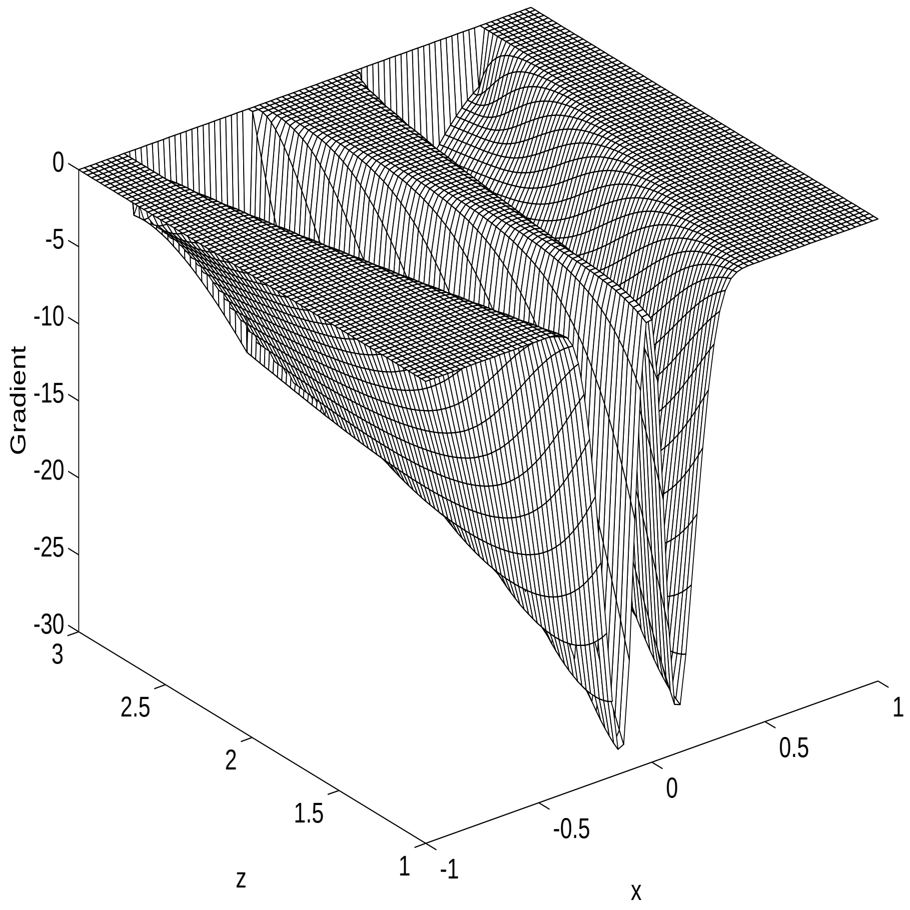

Figure 4.4: L^2 Gradient of the Cost function in homogeneous medium

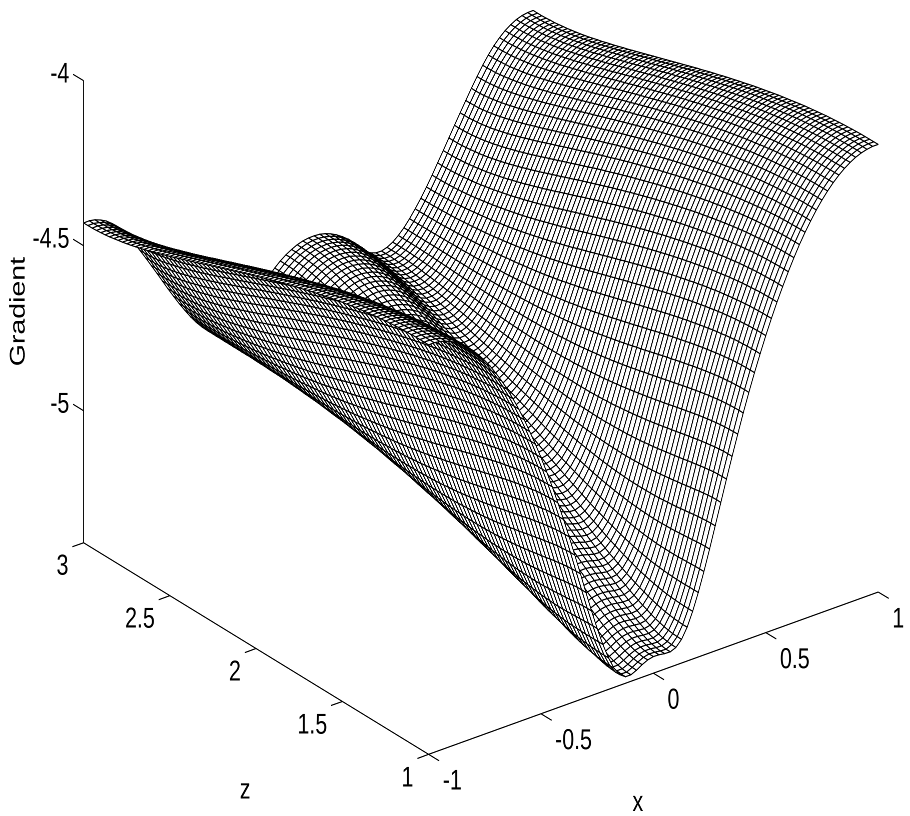

Figure 4.5: H^1 Gradient of the Cost function in homogeneous medium

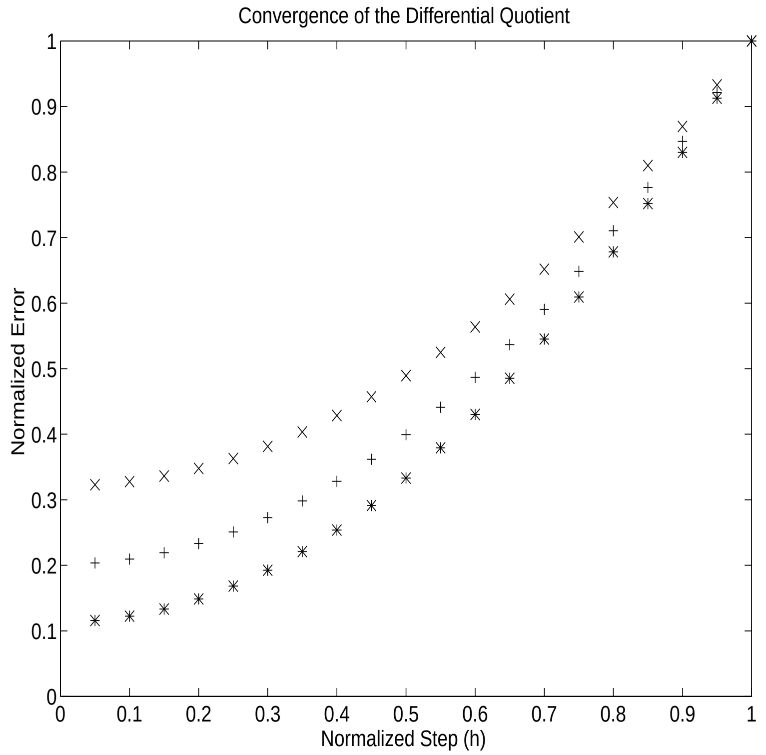

Figure 4.6: Quadratic Convergence of the Differential Quotient to the scalar product of the gradient and the perturbation in the case of homogeneous media

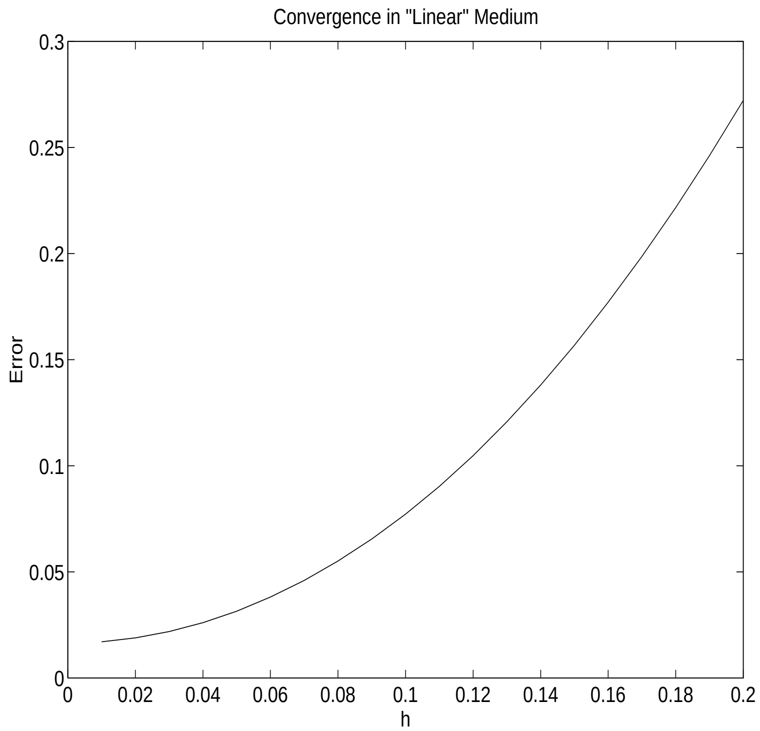

Figure 4.7: Quadratic Convergence of the Differential Quotient to the scalar product of the gradient and the perturbation in the case of linearly increasing velocity