A Note on Consistency and Adjointness for Numerical Schemes

Alain Sei
William Symes

CRPC-TR95527
January 1995

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

A Note on Consistency and Adjointness for Numerical Schemes

Alain Sei and William W Symes *

(Submitted to Mathematical and Computer Modelling)

Abstract

This work deals with the consistency of finite difference approximations. We investigate the relation between the consistency of a numerical scheme and the consistency of its adjoint. We exhibit examples of numerical schemes which are consistent with a (direct) equation and whose adjoint is not consistent with the adjoint equation. This undesirable feature appears in the application of the adjoint state technique which requires an adjointness relation to be satisfied. Therefore, the numerical scheme for the adjoint equation is determined by the choice of the numerical scheme on the direct equation. We conclude that in general consistency is not conserved by adjointness.

Key Words. Finite Differences, Adjoint Schemes, Consistency

AMS(MOS) subject classifications. 65M06, 65M12

1 Introduction

The equivalence theorem (cf [1, 2]) is the fundamental tool to derive convergent finite difference approximations of linear partial differential equations. It states that if a scheme is consistent then stability is equivalent to convergence. Consequently a lot of work has been devoted to stability problems (cf [1, 3, 4]). Indeed the problem of deriving a consistent scheme is generally easily treated by using Taylor's formula. However the problem of consistency can become acute if the choice of the scheme is restricted or even imposed. Such is the case when the adjoint state technique is used.

*Department of Computational and Applied Mathematics, Rice University, P.O Box 1892, Houston, Texas 77251-1892, (sei@rice.edu)

This technique (cf [5, 6, 7]) essentially used in optimization and optimal control gives the gradient of a non linear functional as the solution of a so called adjoint equation. This equation is intrinsic to the functional considered and is defined by satisfying an adjointness relation.
The numerical method used to compute the gradient must also satisfy this adjointness condition. That is why once a numerical scheme has been chosen on the "direct" equation (adjoint of the adjoint), the numerical scheme for the adjoint equation is implicitly determined.
A natural question to ask is whether the numerical approximation for the adjoint equation converges if the numerical approximation for the direct equation does. In other words, is convergence conserved by adjointness ? With the equivalence theorem in mind this question can be rephrased as "Are stability and consistency conserved by adjointness ?".
To answer that question negatively we will exhibit a few simple examples where consistency of the numerical scheme on the direct equation does not imply consistency for the adjoint scheme of the adjoint equation.
The paper is organized as follows. In section 2 we recall briefly the principle of the adjoint state method. In section 3 we apply this technique to the particular problem of traveltime tomography inversion. In section 4 we give a few examples of non consistent approximation for the adjoint equation and we present our conclusions in section 5.

2 Adjoint state technique

The adjoint state technique is commonly used in optimization and optimal control to compute gradients of non linear functionals. This technique consists in computing an auxiliary field (the adjoint state) by solving a so called adjoint equation. This adjoint state enters directly in the computation of the gradient of the functional. For example let's consider the following least squares problem :

Find the Minimizer m^* of the functional

$$J(m) = \frac{1}{2}||F(m) - D||^2_{\mathcal{D}}$$

where F is a non linear operator from a certain Hilbert "model" space M into a Hilbert "data" space D:

$$\begin{array}{rcl} F : \mathcal{M} & \longrightarrow & \mathcal{D} \\ m & \longmapsto & F(m) = d \end{array}$$

To find the minimizer m^* with local optimization techniques, we need to compute the gradient of the functional J. The derivative of J at m in the direction δm is given by:

$$J'(m).\delta m \;=\; (\; F'(m).\delta m, F(m) - D \;)_{\mathcal{D}} \tag{2.1}$$

where $(.,.)_{\mathcal{D}}$ is the scalar product in $\mathcal{D}$. The derivative $F'(m)$ of the functional F at m is a linear operator from $\mathcal{M}$ to $\mathcal{D}$. If we assume that $F'(m)$ is a one to one mapping then it admits an inverse $\mathcal{L} = (F'(m))^{-1}$ defined as follows:

$$\begin{array}{rcl} \mathcal{L} : \mathcal{D} & \longrightarrow & \mathcal{M} \\ \delta d & \longmapsto & \mathcal{L}\,\delta d = \delta m \end{array}$$

Now let's introduce a new field w solution of the adjoint equation:

$$\mathcal{L}^* w = F(m) - D$$

Then equation (2.1) can be written as follows:

$$\begin{array}{rcl} J'(m).\delta m & = & (\; F'(m).\delta m, F(m) - D \;)_{\mathcal{D}} \\ & = & (\; F'(m).\delta m, \mathcal{L}^* w \;)_{\mathcal{D}} \\ & = & (\; \mathcal{L}F'(m).\delta m, w \;)_{\mathcal{M}} \\ & = & (\; \delta m, w \;)_{\mathcal{M}} \end{array}$$

where $(.,.)_{\mathcal{M}}$ is the scalar product in $\mathcal{M}$. Therefore since the gradient G of J is defined as the element of $\mathcal{M}$ such that

$$J'(m).\delta m = (\; \delta m, G \;)_{\mathcal{M}}$$

we have $G = w$.

3 Tomography inversion

We consider as a practical case the inverse tomography problem in seismology. Given travel times data in a domain Ω we want to determine the slowness field m (reciprocal of the velocity) minimizing the misfit between computed traveltimes and traveltimes data. We want to minimize the following functional:

$$J(m) = \frac{1}{2} \int_{\Omega} \left(\tau(x,z;m) - \tau^d(x,z)\right)^2 dx\; dz$$

where $\tau(x,z;m)$ solves the eikonal equation (cf [10, 11]):

$$\left\{\begin{array}{rcll} \tau_x^2+\tau_z^2 & = & m^2 & (x,z)\in\Omega \\ \\ \tau & = & \phi & (x,z)\ \ on\ \Gamma_0 \end{array}\right. \tag{3.1}$$

We assume that the non linear operator F associates to each slowness m a unique travel time τ solution of (3.1). The derivative of J at m in the direction δm is :

$$J'(m).\delta m = \int_\Omega F'(m).\delta m \left(\tau(x,z;m)-\tau^d\right) dx\ dz$$

Let us set $F'(m).\delta m = \delta\tau$. The traveltime perturbation $\delta\tau$ is caused by the perturbation δm in the slowness field. Given δm, $\delta\tau$ is the solution of the following linear equation (cf. Appendix A):

$$\left\{\begin{array}{rcll} \dfrac{\nabla\tau}{m}.\nabla\delta\tau & = & \delta m & (x,z)\in\Omega \\ \\ \delta\tau & = & 0 & (x,z)\ on\ \Gamma_0 \end{array}\right. \tag{3.2}$$

Thus the derivative operator $F'(m)$, defined by $F'(m).\delta m = \delta\tau$ solution of (3.2), is the solution operator of (3.2). Let $\mathcal{L}$ be the inverse operator of $F'(m)$ (we assume it exists), then $\mathcal{L}\ \delta\tau = \delta m$. Therefore, $\mathcal{L}$ is defined by equation (3.2). We define the adjoint state w as the solution of the adjoint equation

$$\mathcal{L}^* w = \tau(m) - \tau^d \tag{3.3}$$

that is :

$$\left\{\begin{array}{rcll} -\nabla.\left(w\dfrac{\nabla\tau}{m}\right) & = & \tau(m)-\tau^d & (x,z)\in\Omega \\ \\ w & = & 0 & (x,z)\ on\ \Gamma_1 = \Gamma-\Gamma_0 \end{array}\right. \tag{3.4}$$

Then we can write

$$J'(m).\delta m \ = \ \int_\Omega \delta\tau\left(\tau(m)-\tau^d\right)\ dx\ dz = \int_\Omega \delta\tau\ \mathcal{L}^* w\ dx\ dz$$

Using the adjointness relation

$$\int_\Omega \delta\tau\left(\mathcal{L}^* w\right)\ dx\ dz \ = \ \int_\Omega \left(\mathcal{L}\delta\tau\right) w\ dx\ dz \tag{3.5}$$

we have

$$J'(m).\delta m \ = \ \int_\Omega \delta m.w\ dx\ dz$$

Therefore w is the L^2 gradient.

4 Numerical Methods

The crucial point of the method is the adjoint relation (3.5), which for differential operators reduces to integration by parts. When we compute the adjoint state, relation (3.5) must also be satisfied by the discrete operators. To solve the adjoint equation (3.3), we approximate the operator $\mathcal{L}^*$ by a discrete operator L_h^*. The discrete equivalent of the adjointness condition (3.5) is satisfied if this operator is the adjoint of an operator L_h which must be an approximation of $\mathcal{L}$.

We will show that some simple schemes do not have this property and therefore that consistency is not conserved by the operation of taking the adjoint.

We illustrate this property on the example introduced in section 2. We want to compute the solution $\delta\tau = u$ of equation (3.2). To simplify the problem, we assume that $\tau_z(x,z;m) \neq 0$ in the domain, that is, there are no turning rays. So we can divide through by τ_z and therefore u is the solution of :

$$(4.1) \qquad \begin{cases} u_z + a(x,z)u_x = f & (x,z) \in \Omega \\ \\ u = 0 & (x,z) \text{ } on \text{ } \Gamma_0 \end{cases}$$

with $a(x,z) = \dfrac{\tau_x}{\tau_z}$ and $f = \dfrac{\delta m}{\tau_z}$. The adjoint equation of (4.1) for the L^2 scalar product is :

$$(4.2) \qquad \begin{cases} -w_z - (a(x,z)w)_x = g & (x,z) \in \Omega \\ \\ w = 0 & (x,z) \text{ } on \text{ } \Gamma_1 = \Gamma - \Gamma_0 \end{cases}$$

These two equations satisfy the following adjointness relation:

$$\int_\Omega (u_z + au_x).w \, dx \, dz = \int_\Omega (-w_z - (aw)_x).u \, dx \, dz$$

We can transform (4.1) and (4.2) into initial value problems (the z variable being considered as time) by a judicious choice of f and g. With $f(x,z) = -a(x,z)\phi'(x)$ equation (4.1) transforms, for the function $\tilde{u} = u + \phi$, into the initial value problem:

$$(4.3) \qquad \begin{cases} \tilde{u}_z + a(x,z)\tilde{u}_x = 0 & (x,z) \in \Omega \\ \\ \tilde{u} = \phi & (x,z) \text{ } on \text{ } \Gamma_0 \end{cases}$$

and with $g(x,z) = (a(x,z)\psi(x))_x$ equation (4.2) transforms, for the function $\tilde{w} = w + \psi$, into the initial value problem:

$$(4.4) \qquad \begin{cases} -\tilde{w}_z - (a(x,z)\tilde{w})_x = 0 & (x,z) \in \Omega \\ \tilde{w} = \psi & (x,z) \; on \; \Gamma_1 = \Gamma - \Gamma_0 \end{cases}$$

Both equation (4.3) and (4.4) are hyperbolic equations. For the particular choice of $a(x,z) = x/z$ we can solve equation (4.3) and (4.4) analytically. So we can compare the numerical solution with the exact solution. In the sequel we fix $a(x,z) = x/z$. We start with the simplest first order upwind scheme.

4.1 Upwind Schemes

4.1.1 First Order Scheme

We consider the following domain $\Omega =]x_0, x_1[\times]z_0, z_1[$, the boundary Γ_0 being the line $z = z_0$. We discretize (4.3) with a first order approximation in z and an upwind derivative in x as follows (cf [8] pp 112):

$$(4.5) \qquad \begin{cases} \dfrac{u_j^{n+1} - u_j^n}{\Delta z} + (a^+)_j^n \dfrac{u_j^n - u_{j-1}^n}{\Delta x} + (a^-)_j^n \dfrac{u_{j+1}^n - u_j^n}{\Delta x} = 0 \\ u_j^1 = \phi_j \end{cases}$$

where $a^+ = max(\dfrac{x}{z}, 0)$, $a^- = min(\dfrac{x}{z}, 0)$. So $(a^+)_j^n = 0$ for $j \le j_0$ and $(a^-)_j^n = 0$ for $j \le j_0$, where j_0 corresponds to $x = 0$.

Equation (4.3) being hyperbolic we expect conditional stability. This scheme is stable if Δz and Δx are chosen such that :

$$(4.6) \qquad \frac{\max|a|\Delta x}{\Delta z} \le 1$$

as is easily seen by a plane wave (or Von Neumann) analysis (cf [3]). The truncation error of this scheme is $O(\Delta z + \Delta x)$ for any point of the (x,z) grid. The adjoint scheme of (4.5) derived in appendix B , is given by:

$$(4.7) \qquad \begin{cases} \dfrac{w_j^{n-1} - w_j^n}{\Delta z} - \dfrac{(a^+)_{j+1}^n w_{j+1}^n - (a^+)_j^n w_j^n}{\Delta x} - \dfrac{(a^-)_j^n w_j^n - (a^-)_{j-1}^n w_{j-1}^n}{\Delta x} = 0 \\ w_1^n = w_J^n = 0 \qquad n = 2..N \qquad w_j^N = \psi_j \qquad j = 1..J \end{cases}$$

This scheme is stable under condition (4.6). Let's examine the truncation in $x = 0$. We have at that point (of index j_0), $(a^+)^n_{j_0} = (a^-)^n_{j_0} = 0$. Therefore we can write:

$$-\frac{(a^+)^n_{j_0+1} w^n_{j_0+1} - (a^+)^n_{j_0} w^n_{j_0}}{\Delta x} - \frac{(a^-)^n_{j_0} w^n_{j_0} - (a^-)^n_{j_0-1} w^n_{j_0-1}}{\Delta x}$$

$$= -\frac{(a^+)^n_{j_0+1} w^n_{j_0+1} - (a^-)^n_{j_0-1} w^n_{j_0-1}}{\Delta x} \;=\; -2.(aw)_x(0,z) + O(\Delta x)$$

The adjoint scheme is not consistent with the adjoint equation. We illustrate this phenomenon in the case where $\Omega =]-1,1[\times]1,3[$ and $\phi(x) = x^2$.

The exact solution of equation (4.4) derived in appendix C is given by

$$w(x,z) = \frac{1}{z}\phi(x\frac{1}{z})$$

We plot below the exact and the numerical solution. The inconsistency of the numerical scheme in 0 is obvious. We can notice that convergence is assured everywhere but at the point $x = 0$. This is in agreement with the equivalence theorem (cf [1, 2]) which implies that if a scheme is stable and is not convergent then it cannot be consistent.

The inconsistency of the adjoint scheme seems to be a direct consequence of the upwind character

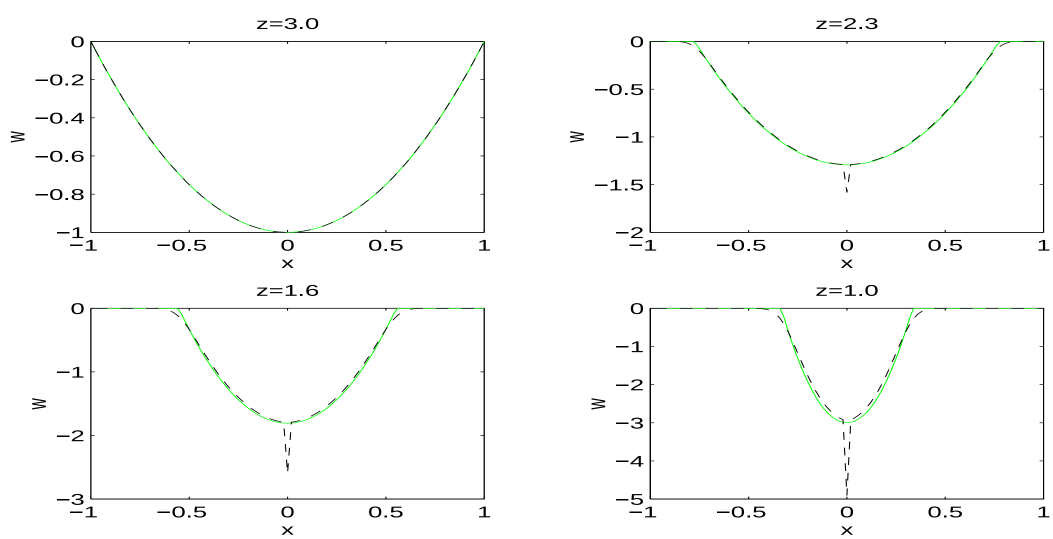

Figure 4.1: The exact solution (solid line) and the numerical solution (dashed line) of the adjoint equation for different depth. The depth $z = 3$ is the initial data curve. The scheme is not consistent in $x = 0$.

of the direct scheme of equation (4.5). The adjoint scheme approximates the following continuous

equation:

$$-w_z - (a^+(x,z)w)_x - (a^-(x,z)w)_x = 0$$

But since a^+ and a^- are defined using the functions Max and Min which are not differentiable in 0, we have inconsistency. This analysis is supported by using another upwind scheme.

4.1.2 Second Order Scheme

We choose to use a second order one sided approximation in space. Since the z-derivative is not a problem we keep a first order approximation in z. So, we consider the following scheme:

$$\text{(4.8)} \qquad \left\{ \begin{array}{rcl} \dfrac{u_j^{n+1} - u_j^n}{\Delta z} + (a^+)_j^n D2_x^- u_j^n + (a^-)_j^n D2_x^+ u_j^n & = & 0 \\ \\ u_j^1 & = & \phi_j \end{array} \right.$$

where $D2_x^-$ (resp $D2_x^+$) is the left (resp. right) second order approximation given by:

$$\left\{ \begin{array}{l} D2_x^- u_j^n = \dfrac{u_{j-2}^n - 4.u_{j-1}^n + 3.u_j^n}{2\Delta x} \\ \\ D2_x^+ u_j^n = -\dfrac{u_{j+2}^n - 4.u_{j+1}^n + 3.u_j^n}{2\Delta x} \end{array} \right.$$

This scheme is consistent with equation (3.4) and the truncation error is $O(\Delta z + \Delta x^2)$ at every point (x,z) of the grid. The adjoint scheme is given by:

$$\text{(4.9)} \qquad \left\{ \begin{array}{l} \dfrac{w_j^{n-1} - w_j^n}{\Delta z} - D2_x^+((a^+)_j^n w_j^n) - D2_x^-((a^-)_j^n w_j^n) = 0 \\ \\ w_1^n = w_J^n = 0 \qquad n = 2..N \qquad w_j^N = \psi_j \end{array} \right.$$

since the adjoint of $D2_x^-$ is $-D2_x^+$ and the adjoint of $D2_x^+$ is $-D2_x^-$. In $x = 0$ we have again $(a^+)_{j_0}^n = (a^-)_{j_0}^n = 0$, and so:

$$D2_x^+((a^+)_j^n w_j^n) + D2_x^-((a^-)_j^n w_j^n) =$$

$$= \frac{(a^+)_{j+2}^n w_{j+2}^n - 4.(a^+)_{j+1}^n w_{j+1}^n + 3.(a^+)_j^n w_j^n}{2\Delta x} - \frac{(a^-)_{j-2}^n w_{j-2}^n - 4.(a^-)_{j-1}^n w_{j-1}^n + 3.(a^+)_j^n w_j^n}{2\Delta x}$$

$$= \frac{(a^+)_{j+2}^n w_{j+2}^n - 4.(a^+)_{j+1}^n w_{j+1}^n}{2\Delta x} - \frac{(a^-)_{j-2}^n w_{j-2}^n - 4.(a^-)_{j-1}^n w_{j-1}^n}{2\Delta x}$$

$$= 2(aw)_x(0,z) - 4(aw)_x(0,z) + O(\Delta x) = -2(aw)_x(0,z) + O(\Delta x)$$

Therefore this scheme is not consistent with equation (4.4) in $x = 0$. Furthermore it is not consistent at $x = \pm\Delta x$. For example at $x = \Delta x$, corresponding to $j = j_0 + 1$ we have $(a^-)^n_{j-1} = (a^-)^n_{j_0} = 0$ and by definition of a^- we have $(a^-)^n_j = (a^-)^n_{j_0+1} = 0$ therefore:

$$D2^+_x((a^+)^n_j w^n_j) + D2^-_x((a^-)^n_j w^n_j) =$$

$$= \frac{(a^+)^n_{j+2}w^n_{j+2} - 4.(a^+)^n_{j+1}w^n_{j+1} + 3.(a^+)^n_j w^n_j}{2\Delta x} - \frac{(a^-)^n_{j-2}w^n_{j-2} - 4.(a^-)^n_{j-1}w^n_{j-1} + 3.(a^+)^n_j w^n_j}{2\Delta x}$$

$$= \frac{(a^+)^n_{j+2}w^n_{j+2} - (a^-)^n_{j-1}w^n_{j-1}}{2\Delta x} - \frac{4.(a^+)^n_{j+1}w^n_{j+1} - 3.(a^+)^n_j w^n_j}{2\Delta x}$$

$$= 2(aw)_x(\Delta x, z) - \frac{4.(a^+)^n_{j+1}w^n_{j+1} - 3.(a^+)^n_j w^n_j}{2\Delta x} + O(\Delta x)$$

as a result, the scheme is not consistent at $x = \Delta x$ (and also $x = -\Delta x$ by symmetry).

The inconsistency of the adjoint scheme seems to be caused by the upwind character of the nu-

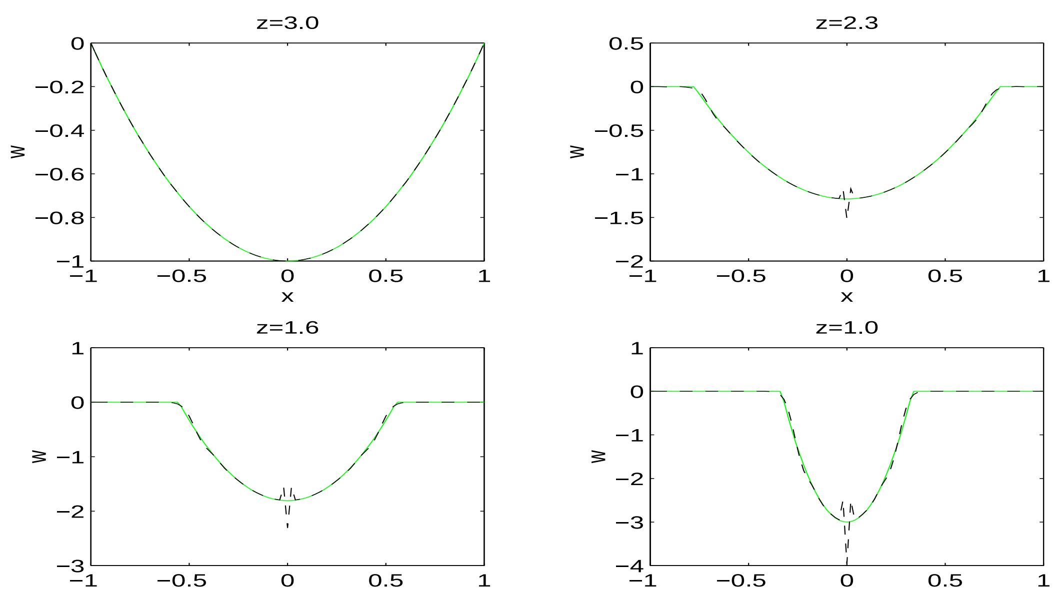

Figure 4.2: The exact solution (solid line) and the numerical solution (dashed line) of the adjoint equation for different depth. The depth $z = 3$ is the initial data curve. The scheme is not consistent in $x = 0$.

merical method chosen. But as we shall see in the next section, problems even occur with centered schemes. We illustrate this point on the Lax-Wendroff scheme (cf [8] pp 101).

4.2 Centered Scheme

The Lax-Wendroff scheme is a centered, dissipative approximation of (4.3) second order in x and z. To get a second order approximation in x we simply use a centered finite difference approximation. To get a second order approximation in z we use the modified equation approach (cf [9, 3]). We have

$$\begin{aligned} u_{zz} &= -(au_x)_z = -a_z u_x - a u_{xz} \\ &= -a_z u_x + a(au_x)_x = -a_z u_x + a^2 u_{xx} + a a_x u_x \end{aligned}$$

We use this expression to derive a second order accurate scheme in the interior of the domain as follows:

$$D_z^+ u_j^n - \frac{\Delta z}{2}\left((a_j^n)^2 \Delta_x u_j^n + a_j^n D_x^o a_j^n D_x^o u_j^n - D_z^+ a_j^n D_x^o u_j^n\right) + a_j^n D_x^o u_j^n = 0 \tag{4.10}$$

where we have used :

$$\begin{cases} D_z^+ u_j^n = \dfrac{u_j^{n+1} - u_j^n}{\Delta z} & D_x^o u_j^n = \dfrac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} \\ \Delta_x = \dfrac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x} & \end{cases}$$

On the boundary we use a first order upwind scheme

$$\begin{cases} D_z^+ u_J^n + (a^+)_J^n \dfrac{u_J^n - u_{J-1}^n}{\Delta x} = 0 \\ D_z^+ u_1^n + (a^-)_1^n \dfrac{u_2^n - u_1^n}{\Delta x} = 0 \end{cases}$$

This scheme is at any point (x, z) second order x and second order in z. Its adjoint is given by taking the adjoint of each operator in (4.10). Using the following relations:

$$\begin{cases} (D_z^+)^* = -(D_z^-) & D_z^- u_j^n = \dfrac{u_j^n - u_j^{n-1}}{\Delta z} \\ (\Delta_x)^* = (\Delta_x) & (D_x^o)^* = -(D_x^o) \end{cases}$$

the adjoint scheme in the interior of the domain is given by:

$$-D_z^- w_j^n - \frac{\Delta z}{2}\left(\Delta_x((a_j^n)^2 w_j^n) - D_x^o(a_j^n D_x^o a_j^n w_j^n) + D_x^o(D_z^+ a_j^n w_j^n)\right) + D_x^o(a_j^n w_j^n) = 0 \tag{4.11}$$

On the boundary we use the adjoint of the first order upwind scheme :

$$\begin{cases} -D_z^- u_J^n - \dfrac{(a^+)_J^n u_J^n - (a^+)_{J-1}^n u_{J-1}^n}{\Delta x} = 0 \\ -D_z^- u_1^n - \dfrac{(a^-)_2^n u_2^n - (a^-)_1^n u_1^n}{\Delta x} = 0 \end{cases}$$

This scheme, unlike the first or second order upwind scheme, is consistent with (3.4). But it is not second order as one would expect. The approximation of the z derivative is given by :

$$-D_z^- w_j^n - \frac{\Delta z}{2}\left(\Delta_x((a_j^n)^2 w_j^n) - D_x^o(a_j^n D_x^o a_j^n w_j^n) + D_x^o(D_z^+ a_j^n w_j^n)\right)$$

which should be a second order approximation of $-w_z$. By Taylor's expansion we have :

$$-w_z = -D_z^- w_j^n - \frac{\Delta z}{2} w_{zz} + O(\Delta z^2)$$

Therefore,

$$Q = \Delta_x((a_j^n)^2 w_j^n) - D_x^o(a_j^n D_x^o a_j^n w_j^n) + D_x^o(D_z^+ a_j^n w_j^n)$$

should be an approximation of w_{zz}. This quantity satisfies :

$$Q = (a^2 w)_{xx} - (a a_x w)_x + (a_z w)_x + O(\Delta x^2)$$

But using equation (4.4) we have

$$w_{zz} = -(aw)_{xz} = -(aw)_{zx} = -(a_z w + a w_z)_x = -(a_z w)_x + (a(aw)_x)_x$$

We can see at once that Q is not an approximation of w_{zz} since for instance the term with a_z has the wrong sign. Therefore since this term is the correction term multiplied by Δz the adjoint scheme is consistent with the adjoint equation. However it cannot be second order as the scheme (4.10).

5 Conclusions

We have shown in this paper that the consistency of a numerical scheme with a continuous equation does not imply the consistency of the adjoint scheme with the adjoint equation. This property should be expected since L_h consistent with L means that the truncation error goes to zero with the mesh size. That is if P_h is the projection from V ot V_h, $L \in \mathcal{L}(V,V)$ and $L_h \in \mathcal{L}(V_h,V_h)$ we have

$$||P_hL - L_hP_h||_{\mathcal{L}(V,V)} \longrightarrow 0 \quad \text{when} \quad h \longrightarrow 0$$

The consistency of the adjoint scheme L_h^* means therefore that

$$||P_hL^* - L_h^*P_h||_{\mathcal{L}(V,V)} \longrightarrow 0 \quad \text{when} \quad h \longrightarrow 0$$

So except when L and L_h are self-adjoint (cf [12]) the consistency of L_h does not imply the consistency of L_h^*. Furthermore when both the chosen scheme and its adjoint are consistent, they do not have necessarily have the same order of accuracy. Therefore the adjoint state technique needs to be applied carefully at the discrete level. In particular the numerical method chosen to compute the adjoint state should have a consistent adjoint.

A Derivation of the perturbed equation

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

$$\left\{\begin{array}{rcll} \nabla\tau(m+\delta m)^2 & = & (m+\delta m)^2 & \quad in \;\; \Omega \\ \\ \tau(m+\delta m) & = & \phi & \quad on \;\; \Gamma_0 \end{array}\right.$$

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

$$\begin{array}{rcl} (\nabla\tau(m+\delta m))^2 & = & (\nabla(\tau(m) + \tau^{'}(m).\delta m + o(\delta m^2)))^2 \\ & = & (\nabla\tau(m))^2 + 2.\nabla\tau(m).\nabla\tau^{'}(m).\delta m + o(\delta m^2)) \end{array}$$

we can write

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

Dropping the term of order greater or equal to two (because we are looking for the first derivative) and dividing by m we find that $\delta\tau$ is the solution of :

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

B Adjoint Upwind Scheme

We are looking for the adjoint equation of equation (4.5) 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, defined by the discrete equation (4.5). 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

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

The second integral can be written as

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

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

Let us treat the last integral

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

Finally we can write

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

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

$$
+ \sum_{n=1}^{N-1}(a^-)_1^n w_1^n u_2^n \Delta z - \sum_{j=1}^{J-1} w_j^{N-1} u_j^N \Delta x - \sum_{n=1}^{N-1}(a^+)_J^n w_J^n u_{J-1}^n \Delta z
$$

Choosing 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 (4.5) is given by (4.7). Using the right handside of equation (4.5) and equation (4.7) 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 Exact solution of equations (4.3) and (4.4)

In the specific case where $a(x,z) = x/z$ equations (4.3) and (4.4) are solvable analytically. We use the method of characteristics (cf [13]). We indicate the solution of the adjoint equation (4.4), the method for equation (4.3) being similar. First we rewrite (4.4)

$$-w_z - (\frac{x}{z} w)_x = 0$$

as follows:

$$-w_z - \frac{x}{z} w_x = \frac{1}{z} w$$

The characteristics of equation (4.4) are the curves $x(t), z(t), w(t)$ solution of the ODE system:

$$\begin{aligned} \frac{dz}{dt} &= -1 & \Longrightarrow \qquad & z(t) = z_1 - t \\ \frac{dx}{dt} &= -\frac{x}{z} & \Longrightarrow \qquad & x(t) = \frac{x_0}{z_1} z(t) \\ \frac{dw}{dt} &= \frac{w}{z} & \Longrightarrow \qquad & w(t) = w_0 \frac{z_1}{z(t)} \end{aligned}$$

Since for $t = 0$ we have $x(0) = x_0$ and $z(0) = z_1$, we have $w_0 = w(x_0, z_1) = \psi(x_0) = \psi(x \frac{z_1}{z})$. Therefore $w(x,z) = \frac{z_1}{z} \psi(x \frac{z_1}{z})$.

References

[1] P.D Lax and R.D Richmyer, *Survey of the Stability of Linear Finite Difference Equations.* Communications on Pure and Applied Mathematics, Vol IX, 1956.

[2] L.V Kantorovich, *Functional Analysis and Applied Mathematics*, Armed Services Technical Information Agency, 1952.

[3] R.D Richmyer and K.W Morton, *Difference Methods for Initial Value Problems.* Interscience Publishers, 1967.

[4] R. Courant, K. Friedrichs and H. Lewy, *On the Partial Difference Equations of Mathematical Physics.* Mathematische Annalen Vol 100, 32-74, 1928. *English Translation in IBM Journal March 1967.*

[5] J.L. Lions, *Contrôle optimal de systemes gouvernés par des équations aux dérivees partielles.* Dunod-Gauthier Villard, Paris, 1968 (English Translation : Springer-Verlag, New-York,1982)

[6] 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.

[7] A. Bamberger, G. Chavent, Ch. Hemon and P. Lailly, *Inversion of normal incidence seismograms*, Geophysics, Vol 47, pp 757-770, 1982.

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

[9] P.D Lax and B. Wendroff, *Difference Schemes for Hyperbolic Equations with High Order of Accuracy.* Communications on Pure and Applied Mathematics, Vol XVII, pp 381-398, 1964.

[10] M. Born and E. Wolf, *Principles of Optics; electromagnetic theory of propagation, interference and diffraction of light*, Macmillan Co, New York, 1964.

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

[12] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Ungar Publishing Co, New York, 1955.

[13] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol 2, Interscience Publishers, New York, 1962.