Far Field Boundary Conditions and Perfectly Matched Layer for Some Wave Propagation Problems

Inaugural-Dissertation
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

Tareq Amro

Münster, Mai 2007

Far Field Boundary Conditions and Perfectly Matched Layer for Some Wave Propagation Problems

Inaugural-Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
im Fachbereich Mathematik und Informatik
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität Münster

eingereicht von

Tareq Amro

betreut durch
Prof. Dr. Anton Arnold

am
Institut für Numerische Mathematik
der Westfälischen Wilhelms-Universität Münster
Einsteinstraße 62, D-48149 Münster

Münster, Mai 2007

Dekan:	Prof. Dr. Dr. h.c. Joachim Cuntz
Erster Gutachter:	Prof. Dr. Anton Arnold
Zweiter Gutachter:	Prof. Dr. Martin Burger
Tag der mündlichen Prüfung:	25. Juni 2007
Tag der Promotion:	13. July 2007

Summary

Wave phenomena are the key features in many fields of application, such as fluid dynamics, electromagnetics, aerodynamics, acoustics, seismics, oceanography, and optics. In these fields, accurate numerical simulation of wave motion is crucial for the understanding of basic phenomena as well as for the design and the development of various engineering applications.

These traveling waves phenomena are described by partial differential equations on large and often unbounded spatial domains. To solve such equations numerically, the simulation should first be confined to bounded subdomains of the original domain – of course including that region, where the most interesting wave phenomena take place. As a result, artificial boundaries and corresponding boundary conditions emerge. Four main methods can be used to truncate problems on unbounded or "large" domains: boundary integral methods, infinite element methods, absorbing boundary condition methods, and absorbing layer methods. In this work, different aspects of absorbing boundary conditions and absorbing layers are considered.

This dissertation is split in two parts, each one consists of two chapters:

I. In the first chapter, we study the far field boundary conditions for first order hyperbolic systems. These boundary conditions combine the properties of absorbing boundary conditions (ABCs) for transient solutions and the properties of transparent boundary conditions for steady state problems. Far field boundary conditions were first presented by Engquist and Halpern, and they defined them up to an arbitrary matrix factor. In this work, we develop a general strategy to specify this matrix factor in an optimal way with respect to the absorption of outgoing waves. This is done first by separating the spurious and physical waves, then by minimizing the reflections of the spurious waves from the boundary. Well-posedness of the resulting initial boundary value problem is studied, and convergence in time of the transient solution to the steady state is established.

In the second chapter, we introduce a finite difference scheme to solve the initial boundary value problem established in the first chapter. We apply the well known stability theory due to Gustafsson, Kreiss, and Sundström (GKS-theory) to prove the stability of this scheme. Numerical examples are given, on one hand, to discuss the convergence (as $t \to \infty$) of 2×2 and 3×3 model systems with first order far field boundary conditions to the correct steady state. On the other hand, to compare the numerical solutions for different choices of the scaling matrix.

II. In the second part, we focus on the perfectly matched layer (PML) method. In the third chapter, we propose a PML approach for the numerical solution to nonlinear Klein-Gordon (KG) equations. The procedure includes four steps: Firstly, the nonlinear KG equation is transformed into a semilinear hyperbolic system with a damping term by introducing auxiliary unknown functions. Secondly, we linearize the damping term and design a PML formulation for the linearized system. Then, we derive a nonlinear PML system by replacing the linearized damping term with its original nonlinear counterpart. Finally, an implicit-explicit finite difference scheme is used to solve the nonlinear PML system. This approach is next extended to the two-dimensional case. The numerical tests show the efficiency of this "PML linearization" over other local ABCs.

In the fourth chapter, we present a new PML formulation for the two-dimensional nonlinear compressible and time-dependent Euler equations of fluid dynamics. Both uniform flow and nonuniform but parallel flow in ducted and open domains are considered. We apply the PML technique to the linearized Euler equations. Then the nonlinear PML equations are formed by replacing the linearized flux functions with their nonlinear counterparts. This formulation has an advantage in the form of its hyperbolic part in which the numerical schemes for nonlinear conservation laws can be used directly. We propose an approach to determine the involved layer parameters in a way to guarantee the damping of all the outgoing wave modes inside the PML layers. Different tests are performed and the results demonstrate the effectiveness of the proposed PML.

Key Words: Absorbing boundary conditions, hyperbolic systems, perfectly matched layer, Klein-Gordon equation, Euler equations.

Acknowledgments

I wish to express my sincere gratitude and deep appreciation to my Ph.D. supervisor Prof. Anton Arnold for his close supervision, patience and support. It has been a pleasure to be a student of such an enthusiastic and skilled researcher.

I would like to thank Prof. Chunxiong Zheng for fruitful discussions and helpful remarks. Many thanks to my colleagues in the research group for their cooperation and friendly work atmosphere.

I am grateful to the Mathematical Institute at Münster University for providing me with an ideal working environment. DAAD is greatly acknowledged for presenting the grant to carry out this work.

Finally, I would like to take this opportunity to thank my family, my wife Catrin and her family for the help and support through the years.

Contents

Chapter 1

Introduction

Several physical phenomena of great importance for applications are described by equations whose solutions are composed of waves. An important part of these problems are posed on unbounded domains. To compute a numerical solution to such problems, it is necessary, due to finite computational resources, to truncate the unbounded domain. This is done by introducing an artificial boundary Γ, defining a new domain Ω, which we will refer to as the computational domain. For the problem to be well-posed, it must be closed with a suitable boundary condition on Γ. Also, special care have to be taken when choosing the boundary condition, so that the solution on Ω will be close to the solution on the unbounded domain.

Often the artificial boundary Γ is placed in the far field where the solution is composed of linear waves traveling out of Ω. The fundamental observation is therefore that all reflections caused by the boundary condition on Γ will contaminate the solution in the interior. Hence, for linear waves, the boundary condition should absorb the energy at the artificial boundary. Right in this context, such a boundary condition is usually called absorbing boundary condition (ABC). Other names are also popularly used in the literature, such as nonreflecting, transparent, and open boundary conditions.

The development of better boundary conditions is important, since it will allow for more accurate simulation of wave phenomena in many areas. One such area is aeroacoustics, which is the enabling science for control of acoustics in early design stages of aircraft, cars, and trains. Design for noise reduction in an aircraft is a typical example of a problem posed on an unbounded domain. Simulation of elastic waves, to predict strong ground motions, earthquakes and other geological phenomena, is

another example.
The level of difficulty of constructing a particular boundary condition is determined by the underlying problem. It is possible to divide the boundary conditions into four different categories based on the underlying problem. In increasing order of difficulty they are

- Linear time-harmonic wave propagation problems,
- Linear constant coefficient time-dependent wave propagation problems,
- Linear variable coefficient time-dependent wave propagation problems,
- Nonlinear time-dependent wave propagation problems.

To date, there are accurate and efficient boundary conditions available for linear time-harmonic problems, see the review articles [**30, 78**], and we do not consider such problems here.
For linear, constant coefficient, time-dependent, wave propagation problems, there are ABCs available, which work well for some specific problems. Examples of such problems are Maxwell's equations and the wave equation, for which the perfectly matched layer method (discussed below) is used today, with satisfactory results. For other problems in this category, e.g. anisotropic elasticity, the available methods have not yet reached a fully mature state.
For linear, variable coefficient, time-dependent, wave propagation problems and for nonlinear, time-dependent, wave propagation problems only primitive methods are available. In order to develop methods for these classes of problems, it is important to first have a good understanding of the properties (such as well-posedness and stability) of the corresponding linear problems.
In this thesis, we mainly consider boundary conditions for linear and nonlinear, constant coefficient, time-dependent, wave propagation problems. Here, we are concerned with boundary conditions which are easy to implement and efficient, with respect to both memory and computational time. Also, thinking of the extension to nonlinear problems, we are interested in boundary conditions with good mathematical properties.
In the first part of this introduction, we focus our presentation on ABCs for the wave equation and first order hyperbolic systems. In the second part we review the PML approach for Maxwell's equations and the linearized Euler equations. Often, by limited modifications, many of the methods discussed below can be applied to

other equations as well.
For further studies of wave propagation problems on unbounded domains, we recommend the detailed review articles [**5, 18, 30, 31, 40, 42, 78, 79**].

1.1 Exact and local absorbing boundary conditions

A boundary condition can be defined as a procedure

$$B_E u = 0, \quad \text{on } \Gamma.$$

The boundary condition is said to be exact, if the restriction to Ω, of the solution u, is identical to the solution on an unbounded domain closed with boundary conditions at infinity. Exact boundary conditions are in general, as will be seen below, non-local both in space and time. The non-locality in time often imply storage of the complete temporal history of the solution on the boundary. Storage requirements has been, and still is, the major drawback of exact ABCs. However, there exist novel methods which reduce storage requirements. In addition, these methods can be used together with fast algorithms, reducing the amount of computations per timestep needed to update the boundary condition. In the following we will present exact and local boundary conditions for some common problems.

Wave equation: We start by considering the construction of exact ABCs for the two dimensional wave equation in Cartesian coordinates. Assume that

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \tag{1.1}$$

is to be solved on the half plane $x \geq 0$.
Exact boundary conditions for the wave equation was considered in the classic paper by Engquist and Majda [**24**]. The construction of B_E in [**24**] uses the fact that any leftgoing solution $u(x, y, t)$ of (1.1) can be represented by a superposition of plane waves traveling to the left. Such plane waves are described by

$$u = a e^{i(\sqrt{\xi^2 - \omega^2} x + \xi t + \omega y)}. \tag{1.2}$$

Here a is the amplitude and (ξ, ω) are the duals of (t, y), satisfying $\xi^2 - \omega^2 > 0$ and $\xi > 0$.
Engquist and Majda conclude that, for fixed (ξ, ω), the condition

$$\left(\frac{\partial}{\partial x} - i\sqrt{\xi^2 - \omega^2} \right) u|_{x=0} = 0,$$

annihilates plane waves described by (1.2). For such plane waves this will be an exact ABC. For a more general wave packet, the exact ABC is obtained by superposition. For the wave equation (1.1), the exact ABC becomes

$$\mathcal{F}\left(\frac{\partial u}{\partial x}\right) - i\sqrt{\xi^2 - \omega^2}\mathcal{F}u = 0, \quad \text{on } x = 0, \tag{1.3}$$

where $\mathcal{F}$ denotes the Fourier transform.

The non-locality of the above boundary condition is clearly manifested through the integral transforms. Inverting the transforms directly is not possible, since the function $\sqrt{\xi^2 - \omega^2}$ does not have an explicit inverse transform. However, Engquist and Majda present a method to localize this exact boundary condition. They conclude that if the function

$$\sqrt{1 - \frac{\omega^2}{\xi^2}}$$

is approximated by some rational function, it is possible to localize the approximation by explicitly inverting the Fourier transforms. As a result a hierarchy of local ABCs of increasing order of approximation are produced. The most natural approximation is perhaps to use a Taylor series expansion, however, the resulting local boundary condition leads to an ill-posed problem and can therefore not be used. Engquist and Majda investigate the boundary conditions obtained if a Padé expansion is used and show that the resulting boundary conditions are well-posed. The boundary conditions obtained from the two first Padé expansions are

$$B_1 u \equiv \left(\frac{\partial}{\partial x} - \frac{1}{c}\frac{\partial}{\partial t}\right) u = 0,$$

$$B_2 u \equiv \left(\frac{1}{c}\frac{\partial^2}{\partial t \partial x} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \frac{1}{2}\frac{\partial^2}{\partial y^2}\right) u = 0.$$

The Padé expansion is not the only possible approximation. Other possible expansions (Chebyshev, least-squares, etc) have been studied by Trefethen and Halpern [**77**], where theorems determining the well-posedness for these types of expansions are also presented.

The use of one-way equations as boundary condition was also studied by Higdon [**48**]. He constructed an asymptotically ($m \to \infty$) exact ABC by factorization of one-way equations. The ABC he proposes is given by

$$B_m u = \left(\prod_{j=1}^{m}\left((\cos\alpha_j)\frac{\partial}{\partial x} - c\frac{\partial}{\partial t}\right)\right) u = 0. \tag{1.4}$$

The boundary condition suggested by Higdon is exact for plane waves hitting the boundary at angles $\pm\alpha_j$. For large m, direct application of (1.4) requires discretization of high order derivatives of u. As a result only boundary conditions with small m where used in [**48**].

Bayliss and Turkel [**11**] work with sequences of local ABCs for the wave equation in spherical and cylindrical coordinates. Their boundary condition is similar in structure to (1.4) and is given by

$$B_m u = \left(\prod_{j=1}^{m} \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial r} + \frac{2l-1}{R} \right) \right) u = 0. \tag{1.5}$$

The suggested boundary condition is extendible to high-order, but only the second order formulation is implemented in [**11**].

Another useful tool, which has been used to derive exact boundary conditions of (1.1) is the Dirichlet to Neumann (DtN) map. The DtN map is an operator relating the Dirichlet datum to the Neumann datum on the boundary Γ, enforcing desired asymptotic behavior of the solution at infinity.

For example, consider solutions of the Helmholtz equation posed on the residual domain Σ

$$s^2 \hat{u} = \nabla^2 \hat{u}, \quad x \in \Sigma, \tag{1.6}$$

with boundary conditions at infinity identical to those used for $\Xi \equiv \Sigma \cup \Omega$. Since no boundary condition has been imposed on Γ, there are infinitely many solutions satisfying the above equation. However, the solution on Ξ must be one of these. The desired solution $\hat{u}$, coinciding with the solution on Σ, can be singled out by a specific choice of the operator $\mathcal{D}$

$$\frac{\partial \hat{u}}{\partial n} = -\hat{\mathcal{D}} \hat{u}, \quad x \in \Gamma.$$

Here the normal is taken outward from Ω. The DtN map, $\mathcal{D}$, can be used to define the exact boundary condition B_E

$$B_E u \equiv \frac{\partial u}{\partial n} + \mathcal{L}^{-1}(\hat{\mathcal{D}} \mathcal{L} u), \tag{1.7}$$

where $\mathcal{L}$ denotes the Laplace transform.

Now, (1.7) can be used to derive the exact boundary condition for (1.1) at a planar boundary. As before we assume that (1.1) is solved for Ω being the half plane $x > 0$. Hence, to derive the DtN map we consider solutions which are bounded on Σ. By

taking the Fourier and Laplace transform (with the duals (k, s)) of (1.1) we obtain the ordinary differential equation

$$\frac{\partial^2 \tilde{\hat{u}}}{\partial x^2} = (s^2 + |k|^2)\tilde{\hat{u}}, \quad x \in \Sigma.$$

For $\Re s > 0$, solutions that are bounded on Σ can be written as

$$ae^{\sqrt{s^2+|k|^2}},$$

and the DtN map is therefore defined by

$$\tilde{\hat{\mathcal{D}}} = \sqrt{s^2 + |k|^2}.$$

Here the the branch of the square root is chosen so that $\tilde{\hat{\mathcal{D}}}$ is analytical and positive for $\Re s > 0$. By inserting $\tilde{\hat{\mathcal{D}}}$ into (1.7) we see that the exact boundary condition is identical to the boundary condition (1.3) for $t > 0$. In [**39**], Hagstrom derives a formulation of (1.3), which only involves the inverse Fourier transform in the y direction and a convolution in time. By rewriting $\tilde{\hat{\mathcal{D}}}$ as

$$\tilde{\hat{\mathcal{D}}} = s + (\sqrt{s^2 + |k|^2} - s),$$

the inverse Laplace transform of the function $\hat{K}(s) = \sqrt{s^2 + |k|^2} - s$

$$K(t) \equiv \frac{J_1(t)}{t} = \frac{1}{\pi}\int_{-1}^{1} \sqrt{1-\rho^2}\cos(\rho t)\, d\rho,$$

can be used to derive

$$\frac{\partial u}{\partial n} + \frac{\partial u}{\partial t} + \mathcal{F}^{-1}(|k|^2 K(|k|t) * \mathcal{F}u) = 0. \tag{1.8}$$

By the use of fast algorithms for the computation of convolutions, see [**46**], together with the fast Fourier transform, (1.8) may be directly imposed. This use of fast methods yield an acceptable number of operations required to update the boundary condition.

Note that there is no explicit need for a two dimensional setting in the above examples. The analysis is identical if Γ is chosen as the hyper plane $x = 0$. The Fourier transform is then to be interpreted as the multidimensional Fourier transform.

First order hyperbolic systems: Here we consider the construction of boundary conditions for strongly hyperbolic systems. Let us first give a short overview of one dimensional hyperbolic systems of the form

$$u_t + A(x,t)u_x + C(x,t)u = f(x,t), \tag{1.9}$$

in the strip $0 \leq x \leq 1,\ t \geq 0$, with initial function

$$u(x,0) = u_0(x), \qquad 0 \leq x \leq 1, \tag{1.10}$$

and some boundary conditions, to be fixed later, at $x = 0,\ x = 1$.
$A(x,t), C(x,t) \in I\!R^{n\times n}$ and $f(x,t), u_0(x) \in I\!R^n$ are assumed to be C^∞-smooth functions with respect to all variables. This system is called hyperbolic if $A(x,t)$ has real eigenvalues at every fixed point $x_0 \in [0,1]$ and $t_0 \geq 0$. In particular, we have the following classification.

Definition 1.1. **[36]** *The system in equation (1.9) is called symmetric hyperbolic if A is a Hermitian matrix at every fixed point $x = x_0,\ t = t_0$. It is called strictly hyperbolic if the eigenvalues are real and distinct, it is called strongly hyperbolic if the eigenvalues are real and there exists a complete system of eigenvectors, and, finally, it is called weakly hyperbolic if the eigenvalues are real.*

Strictly and symmetric hyperbolic systems are subclasses of strongly hyperbolic systems. The initial value problem (1.9)-(1.10) is well-posed for strongly hyperbolic systems and is not well-posed for weakly hyperbolic systems. We assume that the system (1.9) is strongly hyperbolic.

Remark 1.2. *Strong hyperbolicity of the system assumes the existence of a smooth transformation $H = H(x,t)$, the rows of H are the left eigenvectors of A, such that*

$$HAH^{-1} = \Lambda = diag(\lambda_1, ..., \lambda_n),$$

where $\lambda_j = \lambda_j(x,t),\ j = 1, ..., n$, are the eigenvalues of A. If we introduce new variables, called characteristic variables, $v = H(x,t)u(x,t)$, then the system (1.9) is transformed to the characteristic form:

$$v_t + \Lambda(x,t)v_x + \breve{C}(x,t)v = \breve{f}(x,t), \tag{1.11}$$

where $\breve{C} = HCH^{-1} + HH_t^{-1} + HAH_x^{-1}$ and $\breve{f} = Hf$. To simplify the notation, we shall assume that the given system is written already in the characteristic variables, thus $A = \Lambda$ in (1.9).

Assumption 1.3. *The eigenvalues of A at the boundary:*

$$\lambda_j(0,t) \text{ and } \lambda_j(1,t), \quad j = 1, ..., n, \tag{1.12}$$

are assumed to be different from zero. Hence, the boundary is called not characteristic. Also they are assumed to have a constant sign as a function of time; i.e., each function of (1.12) is either < 0 *for all* t*, or* > 0 *for all* t*. However, in the interior* $0 < x < 1$ *the eigenvalues may change sign.*

Notation 1.4. u^+, u^- *are used to assemble the variables* u_j *with* $\lambda_j > 0, \lambda_j < 0$*, respectively, at each boundary points.*

Considering the system (1.9) in the characteristic form, we need boundary conditions to determine $u^+(0)$ and $u^-(1)$. Acceptable boundary conditions are to prescribe the ingoing variables at each boundary

$$u^+(0,t) = g_0(t), \tag{1.13a}$$

$$u^-(1,t) = g_1(t), \tag{1.13b}$$

which can be generalized such that the ingoing variables are described in terms of the outgoing ones

$$u^+(0,t) = S_0(t)u^-(0,t) + g_0(t), \tag{1.14a}$$

$$u^-(1,t) = S_1(t)u^+(1,t) + g_1(t), \tag{1.14b}$$

where $S_0(t)$, $S_1(t)$ are matrices of suitable dimensions.
The initial data should be compatible at $t = 0$. Otherwise, the solution would have discontinuities along the characteristics which start at the corners $(x,t) = (0,0)$ and $(x,t) = (1,0)$. Thus, to avoid difficulties connected with nonsmoothness, we shall assume that:

Assumption 1.5. *The initial data vanishes near the corners.*

For example, If we consider the initial boundary problem (1.9)-(1.10) with the boundary conditions (1.14), then, according to the above assumption, the functions g_0, g_1, u_0, f are assumed to vanish near the corners.

Theorem 1.6. *[*[**62**]*, Thm. 7.6.4] Assume that the boundary is not characteristic, and that the data* u_0, f, g_0, g_1 *are compatible* $t = 0$*. The system (1.9), in characteristic form, for* $0 \le x \le 1,\ 0 \le t \le T$ *with initial and boundary conditions (1.10),(1.14) has a unique solution. The solution is a* C^∞*-function, and for every finite time interval* $0 \le t \le T$ *there is a constant* C_T *independent of* u_0, f, g_0, g_1 *such that*

$$\|u(.,t)\|_2^2 + \int_0^t \left(|u(0,\tau)|^2 + |u(1,\tau)|^2\right) d\tau \\ \le C_T \left[\|u_0\|_2^2 + \int_0^t \left(|g_0(\tau)|^2 + |g_1(\tau)|^2 + \|f(.,\tau)\|_2^2\right) d\tau\right] \tag{1.15}$$

for $0 \leq t \leq T$.

The C^∞-smoothness assumptions for the data are made for simplicity. However, once estimates are derived for this case, more general -less smooth- data can be treated as long as the norms for the data are defined, and one obtains a generalized solution [**62**].

Now, we consider the construction of exact ABCs for a first order, multidimensional, constant coefficient, strongly hyperbolic system. The boundary conditions are imposed at a planar boundary, $x = 0$. In the domain Ω, $(x > 0)$ the system can be written as

$$\frac{\partial u}{\partial t} + A\frac{\partial u}{\partial x} + \sum_j B_j \frac{\partial u}{\partial y_j} = 0, \tag{1.16}$$

where $u \in I\!R^n$, $A, B_j \in I\!R^{n\times n}$. If we further assume that A is invertible (no characteristic boundary) we can employ the usual Fourier and Laplace transform to obtain

$$\frac{\partial \tilde{\hat{u}}}{\partial x} = M\tilde{\hat{u}}, \quad M = -A^{-1}\left(sI + \sum_j ik_j B_j\right). \tag{1.17}$$

Since we consider a strongly hyperbolic problem, we can decompose the solution into left and right-going modes or waves. The right-going waves corresponding to positive eigenvalues of M and the left-going to negative eigenvalues.

To decouple the system into two sets of scalar equations we use the diagonalization

$$QMQ^{-1} = \begin{pmatrix} \lambda^+ & 0 \\ 0 & \lambda^- \end{pmatrix} = \Lambda.$$

Here Q is composed of the the eigenvectors of M arranged in two matrices Q^+, Q^- such that

$$\begin{pmatrix} Q^+ \\ Q^- \end{pmatrix}.$$

The resulting scalar problems are

$$\frac{\partial \tilde{\hat{v}}^-}{\partial x} = \lambda^- \tilde{\hat{u}}^-, \quad \frac{\partial \tilde{\hat{v}}^+}{\partial x} = \lambda^+ \tilde{\hat{u}}^+ \tag{1.18}$$

where $\tilde{\hat{v}} = Q\tilde{\hat{u}}$. An exact ABC, which make sure that no waves enter Ω at $x = 0$ is

$$\tilde{\hat{v}}^+ \equiv B^+ \tilde{\hat{u}} = 0. \tag{1.19}$$

It is important to realize that B^+ can be chosen in many ways. In principle, the only restriction on B^+ is that it should be orthogonal to the matrix Q^-. A natural choice

is therefore $B^+ = Q^+$. However, there are cases when this choice is not suitable. One such case is the linearized Euler equations, for which Giles [**28**] discovered that $B^+ = Q^+$ leads to an ill-posed problem. In [**28**], Giles also suggested an alternative choice of B^+, that leads to a well-posed problem. Other choices, leading to a well-posed problem has been suggested by Hagstrom and Goodrich [**34**] and Rowley and Colonius [**71**].

1.2 Perfectly matched layers

Another approach to terminate the computational domain is to surround the scatterer by a finite width, absorbing layer. In any absorbing layer, the governing equations are modified so that the solutions in the layer decay. For such an approach to be effective, all waves traveling into the layer, independent of frequency or angle of incidence, should be absorbed without reflections. Absorbing layers with these ideal properties are referred to as Perfectly Matched Layers (PMLs). PMLs were first introduced in the context of computational electromagnetics by Bérenger [**13**], and is today used widely by engineers in that area.

PMLs for Maxwell's equations: The major breakthrough, that made absorbing layers competitive, compared to global and local ABCs, was the introduction of the Perfectly Matched Layer by Bérenger [**13**]. Bérenger considered Maxwell's equations in two space dimensions

$$\epsilon_0 \frac{\partial E_x}{\partial t} + \sigma E_x = \frac{\partial H_z}{\partial y}, \tag{1.20}$$

$$\epsilon_0 \frac{\partial E_y}{\partial t} + \sigma E_y = -\frac{\partial H_z}{\partial x}, \tag{1.21}$$

$$\mu_0 \frac{\partial H_z}{\partial t} + \sigma^* H_z = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x}. \tag{1.22}$$

Here E_x and E_y are the electric fields, H_z the magnetic field, ϵ_0 and μ_0 are the free space permittivity and permeability and σ and σ^* the electric and magnetic conductivity. Bérenger realized that by the simple splitting $H_z = H_{zx} + H_{zy}$ additional degrees of freedom, that in turn could be used to guarantee the perfect matching, could be introduced. With this splitting he introduced the perfectly matched layer

as a medium governed by the equations

$$\epsilon_0 \frac{\partial E_x}{\partial t} + \sigma_x E_x = \frac{\partial (H_{zx} + H_{zy})}{\partial y}, \tag{1.23}$$

$$\epsilon_0 \frac{\partial E_y}{\partial t} + \sigma_y E_y = -\frac{\partial (H_{zx} + H_{zy})}{\partial x}, \tag{1.24}$$

$$\mu_0 \frac{\partial H_{zx}}{\partial t} + \sigma_x^* H_{zx} = -\frac{\partial E_y}{\partial x}, \tag{1.25}$$

$$\mu_0 \frac{\partial H_{zy}}{\partial t} + \sigma_y^* H_{zy} = \frac{\partial E_x}{\partial y}. \tag{1.26}$$

For these equations, Bérenger showed that if σ_i and σ_i^* satisfy $\sigma_i^* \epsilon_0 = \sigma_i \mu_0$, then there will be no reflections at the medium-PML interface. Also, waves traveling across the interface into the PML should decay exponentially inside the PML at a rate depending on the magnitude of the conductivity parameters σ_i and σ_i^*.
Soon after the introduction of Bérengers PML, Abarbanel and Gottlieb [**1**] showed that the equations (1.23)-(1.26) were only weakly well-posed. There were concerns that Bérengers weakly well-posed PML could become ill-posed, and that its numerical solution would grow at an arbitrary rate. However, there were no reports of such growth in computations.
The results from [**1**] lead to the development of several strongly hyperbolic (hence, well-posed) PMLs. One of these is the PML by Abarbanel and Gottlieb suggested in [**2**]. They assumed that the behavior of the absorbing layer could be described by lossy Maxwell's equations

$$\frac{\partial E_x}{\partial t} = \frac{\partial H_z}{\partial y} + R_1, \tag{1.27}$$

$$\frac{\partial E_y}{\partial t} = -\frac{\partial H_z}{\partial x} + R_2, \tag{1.28}$$

$$\frac{\partial H_z}{\partial t} = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} + R_3. \tag{1.29}$$

Solutions of (1.27)-(1.29) can be written as

$$\begin{pmatrix} H_z \\ E_x \\ E_y \end{pmatrix} = \begin{pmatrix} 1 \\ \Omega_1(x;\alpha,\beta,\omega) \\ \Omega_2(x;\alpha,\beta,\omega) \end{pmatrix} e^{i\omega(t-\alpha x-\beta y)} e^{-\alpha \int_0^x \sigma(\eta)d\eta}.$$

For such solutions to be perfectly matched and decaying, it is required that the unknown functions $R_1, R_2, R_3, \Omega_1, \Omega_2$ satisfy the constraints:

- R_i, $(i = 1, 2, 3)$ should be independent of the parameters α, β, ω,

- the solution should be continuous across the interface,
- the amplitude of the solution vector in the layer should be monotonically decreasing.

Taking these constraints into account and using the dispersion relation

$$\alpha^2 + \beta^2 = 1, \tag{1.30}$$

Abarbanel and Gottlieb derive the following PML

$$\frac{\partial E_x}{\partial t} = \frac{\partial H_z}{\partial y}, \tag{1.31}$$

$$\frac{\partial E_y}{\partial t} = -\frac{\partial H_z}{\partial x} - 2\sigma E_y - \sigma P, \tag{1.32}$$

$$\frac{\partial H_z}{\partial t} = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} + \sigma' Q, \tag{1.33}$$

$$\frac{\partial P}{\partial t} = \sigma E_y, \tag{1.34}$$

$$\frac{\partial Q}{\partial t} = -\sigma Q - E_y. \tag{1.35}$$

For these equations the question of well-posedness is trivial since the system (1.31)-(1.35) is only a zero order perturbation of (1.20)-(1.22) which is well-posed. Also, the auxiliary variables P and Q only appear as ordinary differential equations and will not alter the well-posedness.

PMLs for linearized Euler equations: Acoustic phenomena are governed by the linearized Euler equations. For a problem with oblique flow in two dimensions these take the form

$$r_t + Cr_x + Dr_y + Er = 0, \tag{1.36}$$

where

$$q = \begin{pmatrix} \rho \\ u \\ v \\ p \end{pmatrix}, \quad A = \begin{pmatrix} M_x & 1 & 0 & 0 \\ 0 & M_x & 0 & 1 \\ 0 & 0 & M_x & 0 \\ 0 & 1 & 0 & M_x \end{pmatrix}, \quad B = \begin{pmatrix} M_y & 0 & 1 & 0 \\ 0 & M_y & 0 & 0 \\ 0 & 0 & M_y & 1 \\ 0 & 0 & 1 & M_y \end{pmatrix}.$$

Here (ρ, u, v, p) are non-dimensionalized perturbations of the density, the velocity in x and y direction and the pressure respectively. M_x and M_y are the Mach number

in the x and y directions. The equations (1.36) support three types of waves; sound, entropy, and vorticity waves. Entropy and vorticity waves are convected downstream with the mean flow while the two soundwaves travel up or downstream.

The perfectly matched layer method has also been applied to the linearized Euler equations. The first split-field PML for the linearized Euler equations was suggested by Hu [**51**]. Hu reported that a low pass filter had to be used inside the absorbing layer to suppress exponential instabilities, also Goodrich and Hagstrom reported similar observations in [**34**]. Tam *et al.* [**75**] concluded, from a perturbation analysis of the dispersion relation of Hu's PML, that it supported unstable acoustic modes at certain wave numbers. Hesthaven [**49**] also analyzed Hu's PML and found that it was only weakly well-posed.

The first well-posed PML for the linearized Euler equations was introduced by Abarbanel, Gottlieb and Hesthaven [**3**] for a uniform flow ($M_x = M, M_y = 0$). To be able to apply the construction used for Maxwell's equations (see (1.31)-(1.35)) they used the variable transformation

$$\xi = x, \quad \eta = \sqrt{1 - M^2}y = \gamma y, \quad \tau = Mx + \gamma^2 t. \tag{1.37}$$

In these new variables the dispersion relation, of the transformed equations, is very similar to that of the Maxwell's equations and the techniques used in [**2**] can be applied directly.

Although the PML in [**3**] was well-posed, it still supported exponentially growing modes. To stabilize these modes Abarbanel *et al.* add lower order terms. By using the variable transform (1.37), Hu [**52**] constructed an un-split and stable PML for uniform flow. Hu's PML is well-posed for layers parallel to the flow but only weakly well-posed in corners and in layers perpendicular to the flow. Later, he considered the non-uniform flow case, and proposed a numerical procedure to find such a transformation [**54**]. Recently, Hu extended his idea to the the nonlinear Euler equations [**55**].

A well-posed and stable PML for the linearized Euler equations for a oblique flow has been suggested by Hagstrom in [**42**]. The PML is derived as an example on an application of Hagstrom's general method for the construction of perfectly matched layers for hyperbolic systems.

In his method, Hagstrom considers an absorbing layer of width L at the plane $x = 0$

with the governing equations (for $x < 0$)

$$\frac{\partial u}{\partial t} + A(y)\frac{\partial u}{\partial x} + \sum_j B_j(y)\frac{\partial u}{\partial y_j} + C(y)u = 0, \quad -L < x < 0. \tag{1.38}$$

By Laplace transform in time

$$\hat{u} = e^{\lambda x}\phi, \quad \left(sI + \lambda A + \sum_j B_j(y)\frac{\partial}{\partial y_j} + C(y)\right)\phi = 0, \tag{1.39}$$

it is possible to find the modal solutions to (1.38).
Hagstrom concludes that the solution inside the PML should be modified so that $\Re\lambda$ is bounded away from zero for $\Re s \geq 0$. This condition is equivalent to the variable transform (1.37). By construction the eigenfunctions ϕ are the same at the interface $x = 0$, hence the absorbing layer will be perfectly matched.
Postulating a modal solution inside the PML

$$e^{\lambda x + (\lambda\hat{R}^{-1} - \hat{M}^{-1}\hat{N})\int_0^x \sigma(z)dz}\phi, \tag{1.40}$$

and substituting (1.40) into (1.39) gives the resulting PML equations as

$$\left(sI + \lambda A(I - \sigma(\hat{R} + \sigma)^{-1})\left(\frac{\partial}{\partial x} + \sigma\hat{M}^{-1}\hat{N}\right) + \sum_j B_j(y)\frac{\partial}{\partial y_j} + C(y)\right)\hat{u} = 0.$$

In principle, the operators M, N and R can be very complicated but for most practical applications M and n can be chosen as real scalars and R to be a first order, scalar differential operator in time and in the transverse variables

$$R = \frac{\partial}{\partial t} + \sum_j \beta_j(y)\frac{\partial}{\partial y_j} + \alpha. \tag{1.41}$$

An extension of this PML model (from [**42**]) was presented by Appelö *et al.* [**7**]. In [**83**], we present a PML formulation for the nonlinear Euler equations. This new formulation combine properties of Hagstrom's formulation in [**42**] and that of Hu in [**55**]. In case of non-uniform parallel flow, an approach have been proposed to determine the involved parameters in a way to damp all wave modes in their trajectories in the PML layers.

Notation

- A^T = transpose of the matrix A.
- For $u, v \in I\!R^n$, $\quad \langle u, v \rangle \ = \sum_{i=1}^{n} u_i v_i$,

  for $u, v \in \mathbb{C}^n$, $\quad \langle u, v \rangle \ = \sum_{i=1}^{n} \bar{u}_i v_i$, $\quad |u|^2 = \langle u, u \rangle$
- If A is real symmetric $n \times n$ matrix, we write

$$A \geq \delta I$$

  if $\langle x, Ax \rangle \geq \delta |x|^2$ for all $x \in I\!R^n$
- The space $L^2((0,1); I\!R^n)$, with the norm denoted by $\|.\|_2$, consists of those functions $u : (0,1) \to I\!R^n$, $u = (u_1, ..., u_m)$, with $u_i \in L^2(0,1)$, $\quad (i = 1, ..., n)$.

Part I

Far Field Boundary Conditions

Chapter 2

Far field boundary conditions for linear hyperbolic systems

For the computation of a numerical solution to hyperbolic partial differential equation on an infinite domain, it is common to perform the calculation on a truncated finite domain Ω. This raises the problem of choosing appropriate boundary conditions for the resulting artificial boundary Γ. Ideally, these boundary conditions should prevent any nonphysical reflection of outgoing waves and should be easy to implement numerically. They should also present together with the governing equation a well-posed truncated problem which is a basic requirement for the corresponding numerical approximation to be stable.

Example of hyperbolic equations include the Euler equations of gas dynamics, the shallow water equations, Maxwell's equations, and equations of magnetohydrodynamics. For these hyperbolic problems the correct boundary condition is that waves traveling across the boundary should not be reflected back. Boundary conditions with this property are often referred to as non-reflecting, transparent, artificial or absorbing boundary conditions (ABCs).

The theoretical basis for ABCs stems from a paper by Engquist and Majda [**24**] which discusses both ideal ABCs and a method for constructing approximate forms. In addition, a paper by Kreiss [**61**] which analyzes the well-posedness of the initial boundary value problems for hyperbolic systems. Many researchers have been active in this area in the last years, The readers are referred to [**29, 40, 78, 41, 18, 79**] for further details. However, their work has been mainly concerned with ABCs that are better suited for a transient solution than for a steady solution, and most of

these boundary conditions lead to steady solutions of poor accuracy.
In this chapter, we are concerned with ABCs that lead to accurate steady solutions. In this context, Bayliss and Turkel [**12**] derived non-reflecting conditions for the Euler equations which they used for steady state calculations. These boundary conditions are obtained from expansions of the solution at large distances. Accurate boundary conditions for the steady Euler equations in a channel were also studied by Giles in [**28**].
Our main reference is the paper of Engquist and Halpern [**23**]. They constructed a new class of boundary conditions that combine the properties of ABCs for transient solutions and the properties of transparent boundary conditions for steady state problems. These boundary conditions, which called far field boundary conditions (FBCs), can be used in both the transient regime and when the solution approaches the steady state. In this sense, they can be applied when the evanescent and traveling waves are present in the time-dependent calculation or when a time-dependent formulation is used for computations to steady state. In case of hyperbolic systems, these FBCs are defined up to matrix factor in front of the steady terms [**23**]. This poses the following computational problem (which is one of the main subjects of this chapter): How to choose this factor in a way to accelerate the convergence to the steady state, and to improve the accuracy of the transient solution.
We observe that the FBCs simulate the radiation of energy out of Ω. An incorrect specification of these boundary conditions can cause spurious reflected waves to be generated at Γ. These waves represent energy propagating into Ω. Since they are not part of the desired solution, they can substantially reduce the accuracy of the computed solution. On the other hand, if the time-dependent equations are only an intermediate step toward computing a steady state, then a flow of energy into Ω can delay the convergence to the steady state. Conversely, the correct specification of FBCs can accelerate the convergence. Thus, an answer of the above question consists in minimizing the spurious reflections.
The rest of the chapter is organized as follows: In Section 2, we study briefly the procedure of constructing FBCs for linear hyperbolic systems, and propose a general tool of scaling the included factors. In Section 3, well-posedness and regularity of the resulting initial boundary value problem are studied. General result of convergence in time to the steady state are established in Section 4.

2.1 Derivation of far field boundary conditions

We consider in this section the derivation of hierarchy of FBCs, at $x = 0$ and $x = 1$, for a strictly hyperbolic system of the form

$$u_t + \Lambda u_x + Cu = f(x), \qquad x \in I\!R, \quad t > 0, \tag{2.1}$$

with the initial function

$$u(x, 0) = u_0(x), \qquad x \in I\!R. \tag{2.2}$$

Here C and Λ are $n \times n$ constant matrices, u is a vector with n components. $f(x)$ and $u_0(x)$ are assumed to be C^∞-smooth functions and have supports in $(0, 1)$. The eigenvalues of Λ are distinct and different form zero, that is

$$\Lambda = \begin{pmatrix} \Lambda^+ & 0 \\ 0 & \Lambda^- \end{pmatrix}, \tag{2.3}$$

with $\Lambda^+ = \text{diag}(\lambda_1, ..., \lambda_m), \quad \lambda_j > 0, \qquad \Lambda^- = \text{diag}(\lambda_{m+1}, ..., \lambda_n), \quad \lambda_j < 0.$
We assume that

$$C_1 = \frac{C^T + C}{2} \geq \delta I, \quad \delta > 0, \tag{2.4}$$

which is a necessary condition to ensure the convergence of the whole space problem to the steady state as $t \to \infty$. We will further assume that $\Lambda^{-1}C$ has distinct eigenvalues.

Notation 2.1. *Any $n \times n$-matrix X is partitioned as*

$$X = \begin{pmatrix} X^{++} & X^{+-} \\ X^{-+} & X^{--} \end{pmatrix},$$

where X^{++}, X^{+-}, X^{-+} are $m \times m, m \times (n-m), (n-m) \times m$-matrices, respectively. Also $X^+ := \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} X$ and $X^- := \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} X$.

Solutions of (2.1) are made up of n different modes, which propagate at different speeds. A crucial step on developing boundary conditions for (2.1) is determining the direction of propagation of each mode, and distinguishing which modes are outgoing and which are incoming at the boundary.
If we take a Laplace transform in t, with the dual variable s

$$\tilde{u}(x, s) = \int_0^\infty e^{-st} u(x, t) dt, \quad s \in \mathbb{C}, \quad \Re s > 0,$$

the system becomes

$$\tilde{u}_x + \Lambda^{-1}(sI + C)\tilde{u} = \Lambda^{-1}\tilde{f}.$$

Define $E(s) := \Lambda^{-1} + \frac{1}{s}\Lambda^{-1}C$, we may write

$$\tilde{u}_x + sE(s)\tilde{u} = \Lambda^{-1}\tilde{f}. \tag{2.5}$$

We wish to separate $\tilde{u}$ into "rightgoing" and "leftgoing" modes. Each of these modes corresponds to an eigenvalue of $E(s)$.

Definition 2.2. **[14]** *The inertia of a matrix M is the ordered triple $i(M) = (i_+(M), i_-(M), i_0(M))$, where $i_+(M), i_-(M)$, and $i_0(M)$ are the numbers of eigenvalues of M with respectively positive, negative, and zero real part, all counting multiplicity.*

Lemma 2.3. **[14]** *Let G, H be $n\times n$-matrices with H Hermitian and regular, suppose $HG + G^*H$ is a positive semi-definite and $i_0(G) = 0$. Then $i(G) = i(H)$.*

Lemma 2.4. *For $\Re s > 0$, $E(s)$ has exactly m eigenvalues with positive real part and $(n - m)$ with negative real part; i.e., $i(\Lambda) = i(E)$.*

Proof. Apply Lemma 2.3 with $H := \Lambda$ and $G := \Lambda^{-1}(sI + C)$:

$$\Lambda(\Lambda^{-1}(sI + C)) + (\Lambda^{-1}(sI + C))^*\Lambda = 2\Re sI + 2C_1 > \delta I > 0.$$

Also $i_0(\Lambda^{-1}(sI + C)) = 0$; otherwise $\Lambda^{-1}(sI + C)$ will have a purely imaginary eigenvalue, say $i\omega$, $\omega \in I\!R$. Let ϕ denote its eigenvector. Then

$$i\omega\phi = \Lambda^{-1}(sI + C)\phi \Leftrightarrow (i\omega\Lambda - C)\phi = s\phi,$$

which is impossible since

$$\begin{aligned} 0 &< 2\Re s\,|\phi|^2 = \langle s\phi, \phi\rangle + \langle \phi, s\phi\rangle = \langle (i\omega\Lambda - C)\phi, \phi\rangle + \langle \phi, (i\omega\Lambda - C)\phi\rangle \\ &= -\left\langle \phi, (C + C^T)\phi\right\rangle \leq -2\delta\,|\phi|^2 < 0. \end{aligned}$$

□

From Lemma 2.4 and **[36]**, there is $\eta_0 > 0$ and a nonsingular transformation $T(s)$ such that for $\Re s > \eta_0$,

$$XEX^{-1} = D = \begin{pmatrix} D^+ & 0 \\ 0 & D^- \end{pmatrix}, \tag{2.6}$$

where $D(s)$ is the matrix of eigenvalues of $E(s)$, arranged so that $D^+(s)$ is an $m \times m$ positive definite matrix, corresponding to rightgoing solutions, and $D^-(s)$ is an $(n-m) \times (n-m)$ negative definite matrix, corresponding to leftgoing solutions. Here, we drop the explicit s-dependence; henceforth all the matrices are functions of s unless otherwise noted.

With the characteristic variables $\tilde{v} := X\tilde{u}$ the system (2.5) can be written as

$$\tilde{v}_x + sD\tilde{v} = X\Lambda^{-1}\tilde{f},$$

and then partitioned into

$$\frac{d}{dx}\begin{pmatrix} \tilde{v}^+ \\ \tilde{v}^- \end{pmatrix} - s\begin{pmatrix} D^+ & 0 \\ 0 & D^- \end{pmatrix}\begin{pmatrix} \tilde{v}^+ \\ \tilde{v}^- \end{pmatrix} = X\Lambda^{-1}\tilde{f},$$

where $\tilde{v}^+$ and $\tilde{v}^-$ represent purely "rightgoing" and "leftgoing" modes respectively.

Now, we restrict the domain of x in (2.1) to $(0,1)$. The exact nonreflecting boundary conditions follow immediately. Since there are no incoming modes at a nonreflecting boundary, at the left boundary $x = 0$ there should be no rightgoing modes, so an exact perfectly ABC is

$$\tilde{v}^+ = [X\tilde{u}]^+ = 0, \qquad x = 0. \tag{2.7a}$$

At right boundary, there should be no leftgoing modes, so an exact perfectly ABC is

$$\tilde{v}^- = [X\tilde{u}]^- = 0, \qquad x = 1. \tag{2.7b}$$

Two difficulties arise in implementing the above boundary conditions. First, since the boundary condition is expressed in Laplace transform (x,s)-space, and the matrix $X(s)$ contains non-rational functions of s (e.g., square roots), when we transform back to physical (x,t)-space, the boundary conditions will be nonlocal in time.

From a computational perspective, we would prefer a local boundary condition, which may be obtained by approximating non-rational elements of X by rational functions of s.

The second difficulty arises when approximations are introduced: then the resulting IBVP may be ill-posed. The theory of well-posedness will be discussed in the next section.

For $s \to \infty$, we have $E(s) \to \Lambda^{-1}$ and hence $X(s) \to I$. Following standard practice in [**23**] we shall hence make a high frequency expansion of X for $\Re s > \eta_0$:

$$X(s) = I + \frac{1}{s}X_1 + \frac{1}{s^2}X_2 + O\left(\frac{1}{|s|^3}\right). \tag{2.8}$$

The zero order ABCs are then

$$\tilde{u}^+ = 0, \qquad x = 0,$$
$$\tilde{u}^- = 0, \qquad x = 1.$$

More accurate conditions are obtained by using higher order approximations. First and second order ABCs are respectively

$$\left[(I + \frac{1}{s}X_1)\tilde{u}\right]^+ = 0, \qquad x = 0,$$
$$\left[(I + \frac{1}{s}X_1)\tilde{u}\right]^- = 0, \qquad x = 1,$$

$$\left[(I + \frac{1}{s}X_1 + \frac{1}{s^2}X_2)\tilde{u}\right]^+ = 0, \qquad x = 0,$$
$$\left[(I + \frac{1}{s}X_1 + \frac{1}{s^2}X_2)\tilde{u}\right]^- = 0, \qquad x = 1,$$

which is transformed to the physical space by the substitution $s \to \frac{\partial}{\partial t}$.
Hence, the first and second order ABCs are respectively [**24**]

$$\left[(\frac{\partial}{\partial t} + X_1)u\right]^+ = 0, \qquad x = 0, \tag{2.9a}$$
$$\left[(\frac{\partial}{\partial t} + X_1)u\right]^- = 0, \qquad x = 1, \tag{2.9b}$$

$$\left[(\frac{\partial^2}{\partial t^2} + X_1\frac{\partial}{\partial t} + X_2)u\right]^+ = 0, \qquad x = 0, \tag{2.10a}$$
$$\left[(\frac{\partial^2}{\partial t^2} + X_1\frac{\partial}{\partial t} + X_2)u\right]^- = 0, \qquad x = 1. \tag{2.10b}$$

For large $\Re s > \eta_0$, the term $\frac{1}{s}\Lambda^{-1}C$ in $E(s)$ is a perturbation of Λ^{-1}. In this case, D in (2.6) can be considered as a diagonal matrix [**50**]. With a high frequency expansion D is written as

$$D(s) = \Lambda^{-1} + \frac{1}{s}D_1 + \frac{1}{s^2}D_2 + O\left(\frac{1}{|s|^3}\right),$$

where $D_j(s),\ j = 1, 2, ...$ are diagonal.
Write (2.6) as $XE = DX$, then the $O\left(|s|^{-1}\right)$-equation reads

$$X_1\Lambda^{-1} + \Lambda^{-1}C = \Lambda^{-1}X_1 + D_1.$$

Solving for X_1 and D_1 gives

$$D_1 = diag\left(\frac{c_{11}}{\lambda_1}, ..., \frac{c_{nn}}{\lambda_n}\right),$$

and

$$(X_1)_{jk} = \begin{cases} 0, & j = k, \\ \dfrac{\lambda_k c_{jk}}{\lambda_k - \lambda_j}, & j \neq k, \end{cases}$$

where c_{jk} is the $(j,k)^{th}$ entry of C. The second order expansion of (2.6) yields

$$X_2\Lambda^{-1} + X_1\Lambda^{-1}C = \Lambda^{-1}X_2 + D_1X_1 + D_2.$$

Solving for X_2 and D_2 yields

$$D_2 = diag\left(\sum_{k\neq 1}\frac{c_{1k}c_{k1}}{\lambda_k - \lambda_1}, ..., \sum_{k\neq n}\frac{c_{nk}c_{kn}}{\lambda_k - \lambda_n}\right),$$

and

$$(X_2)_{jk} = \begin{cases} 0, & j = k, \\ \dfrac{1}{\lambda_j - \lambda_k}\left[c_{jj}c_{jk}\dfrac{\lambda_k^2}{\lambda_k - \lambda_j} - \sum_{l\neq j} c_{jl}c_{lk}\dfrac{\lambda_j\lambda_k}{\lambda_l - \lambda_j}\right], & j \neq k. \end{cases}$$

Let us now turn to the stationary problem corresponding to (2.1) :

$$u_x + \Lambda^{-1}Cu = \Lambda^{-1}f(x), \qquad x \in I\!R, \tag{2.11}$$

$$u(x) \to 0, \qquad x \to \pm\infty. \tag{2.12}$$

The following lemma is similar to Lemma 2.4 but for the case $s = 0$.

Lemma 2.5. $i(\Lambda) = i(\Lambda^{-1}C)$.

Proof. Apply Lemma 2.3 with $H := \Lambda$ and $G := \Lambda^{-1}C$

$$\Lambda(\Lambda^{-1}C) + (\Lambda^{-1}C)^T\Lambda = 2C_1 > 0.$$

Assume that $\Lambda^{-1}C$ has the purely imaginary eigenvalue $i\omega$. Then

$$i\omega\phi = \Lambda^{-1}C\phi \Leftrightarrow i\omega\Lambda\phi = C\phi.$$

But, on the other hand

$$\begin{aligned} 0 &= \langle i\omega\Lambda\phi, \phi\rangle + \langle \phi, i\omega\Lambda\phi\rangle \\ &= \langle C\phi, \phi\rangle + \langle \phi, C\phi\rangle = \left\langle \phi, (C + C^T)\phi\right\rangle \geq 2\delta\left|\phi\right|^2 > 0, \end{aligned}$$

and hence $i_0(\Lambda^{-1}C) = 0$. □

Using Lemma 2.5 and that $\Lambda^{-1}C$ has distinct eigenvalues, we diagonalize the system (2.11)

$$w_x + Rw = S\Lambda^{-1}f(x), \tag{2.13}$$

where w is given by $w := Su$,

$$S\Lambda^{-1}CS^{-1} = R = \begin{pmatrix} R^+ & 0 \\ 0 & R^- \end{pmatrix}, \tag{2.14}$$

$R^+ = \text{diag}(r_1, ..., r_m), \ \Re r_j > 0, \qquad R^- = \text{diag}(r_{m+1}, ..., r_n), \ \Re r_j < 0.$
The following boundary conditions for the steady problem on the bounded domain $(0, 1)$ are satisfied by the steady solution on the unbounded domain

$$(Su)^+ = 0, \qquad x = 0, \tag{2.15a}$$

$$(Su)^- = 0, \qquad x = 1. \tag{2.15b}$$

This is true since the general solution of (2.13) outside the support of f is

$$w(x) = \begin{pmatrix} w^+(0)e^{-R^+x} \\ \\ w^-(0)e^{-R^-x} \end{pmatrix}, \qquad x \le 0, \qquad\qquad w(x) = \begin{pmatrix} w^+(1)e^{R^+(1-x)} \\ \\ w^-(1)e^{R^-(1-x)} \end{pmatrix}, \qquad x \ge 1,$$

where $w = (w^+, w^-)^T$ is partitioned in the same way as u. For the decay condition (2.12) to be valid, it is necessary that

$$w^+(0) = w^-(1) = 0.$$

(2.15) is unique up to a multiplication by regular matrices V^+ and V^-, respectively

$$(VSu)^+ = 0, \qquad x = 0, \tag{2.16a}$$

$$(VSu)^- = 0, \qquad x = 1. \tag{2.16b}$$

In [**23**] the authors defined a family of first order FBCs from a combination of the first order ABCs (2.9) and the transparent steady boundary conditions (2.16):

$$\left[(\frac{\partial}{\partial t} + VS)u\right]^+ = 0, \qquad x = 0, \tag{2.17a}$$

$$\left[(\frac{\partial}{\partial t} + VS)u\right]^- = 0, \qquad x = 1, \tag{2.17b}$$

which is defined up to a matrix factor, $V = \begin{pmatrix} V^+ & 0 \\ 0 & V^- \end{pmatrix}$, in front of S. Higher order boundary conditions can formally be derived analogously

$$\left[(\frac{\partial^2}{\partial t^2} + X_1\frac{\partial}{\partial t} + VS)u\right]^+ = 0, \qquad x = 0, \tag{2.18a}$$

$$\left[(\frac{\partial^2}{\partial t^2} + X_1\frac{\partial}{\partial t} + VS)u\right]^- = 0, \qquad x = 1. \tag{2.18b}$$

The solution of the IVP (2.1) on $(0,1)$ with the boundary conditions (2.17) or (2.18), for arbitrary regular V, converges for long time to the steady state, see [**23**] and Section 4. But since spurious reflections pollute the computed solution, a good choice of V^+ and V^- that annihilate the spurious reflections up to higher order can accelerate this convergence for long time computations and gives higher accuracy for short time computations.

To clarify that, we transform the first order left boundary condition (2.17a) into Laplace space, and use the Notation 2.1

$$\left[(I + \frac{1}{s}VS)\tilde{u}\right]^+ = \left(I^+ + \frac{1}{s}V^+S^{++} \;\;,\;\; \frac{1}{s}V^+S^{+-}\right)\tilde{u} = 0. \tag{2.19}$$

In terms of the characteristic variables, $\tilde{u} = X^{-1}\tilde{v}$, where

$$X^{-1}(s) = I - \frac{1}{s}X_1 - \frac{1}{s^2}(X_2 - X_1^2) + O\left(\frac{1}{|s|^3}\right).$$

(2.19) then becomes

$$\left(I^+ + \frac{1}{s}V^+S^{++} \;\;,\;\; \frac{1}{s}V^+S^{+-}\right)\begin{pmatrix} I^+ - \frac{1}{s}X_1^{++} & -\frac{1}{s}X_1^{+-} \\ -\frac{1}{s}X_1^{-+} & I^- - \frac{1}{s}X_1^{--} \end{pmatrix}\tilde{v} + O\left(\frac{1}{|s|^2}\right)$$

$$= \left[I^+ + \frac{1}{s}(V^+S^{++} - X_1^{++})\right]\tilde{v}^+ + \frac{1}{s}\left[V^+S^{+-} - X_1^{+-}\right]\tilde{v}^- + O\left(\frac{1}{|s|^2}\right) = 0.$$

Neglecting $O(|s|^{-2})$-terms, we may solve for the incoming (rightgoing) modes in terms of outgoing ones as long as $\left[I^+ + \frac{1}{s}(V^+S^{++} - X_1^{++})\right]$ is nonsingular (this holds true at least for $|s|$ large)

$$\tilde{v}^+(0) = -\left[sI^+ + (V^+S^{++} - X_1^{++})\right]^{-1}\left[V^+S^{+-} - X_1^{+-}\right]\tilde{v}^-(0) =: R_c^+\tilde{v}^-(0),$$

where R_c^+ is the matrix of reflection coefficients.
Similarly the right boundary condition (2.17b) may be written in term of the characteristic variables as

$$\left(\frac{1}{s}V^-S^{-+} \quad , \quad I^- + \frac{1}{s}V^-S^{--}\right)\begin{pmatrix} I^+ - \frac{1}{s}X_1^{++} & -\frac{1}{s}X_1^{+-} \\ -\frac{1}{s}X_1^{-+} & I^- - \frac{1}{s}X_1^{--} \end{pmatrix}\tilde{v}$$
$$= \frac{1}{s}\left[V^-S^{-+} - X_1^{-+}\right]\tilde{v}^+ + \left[I^- + \frac{1}{s}(V^-S^{--} - X_1^{--})\right]\tilde{v}^- + O\left(\frac{1}{|s|^2}\right) = 0.$$

Neglecting $O(|s|^{-2})$-terms and solving for the incoming (leftgoing) modes in terms of outgoing ones as long as $\left[I^- + \frac{1}{s}(V^-S^{--} - X_1^{--})\right]$ is nonsingular, gives

$$\tilde{v}^-(1) = -\left[sI^- + (V^-S^{--} - X_1^{--})\right]^{-1}\left[V^-S^{-+} - X_1^{-+}\right]\tilde{v}^+(1) =: R_c^-\tilde{v}^+(1),$$

where R_c^- is the matrix of reflection coefficients at the right boundary.
For the pair of boundary conditions to be absorbing up to order $O(|s|^{-2})$, the matrices R_c^+ and R_c^- must be identically zero, that is $V^+S^{+-} - X_1^{+-}$ and $V^-S^{-+} - X_1^{-+}$ must be zero. So the optimal choices of V^+ and V^- are then given as solutions of

$$V^+S^{+-} = X_1^{+-}, \tag{2.20a}$$
$$V^-S^{-+} = X_1^{-+}. \tag{2.20b}$$

If $(S^{+-})^{-1}$ exists, then $V^+ = X_1^{+-}(S^{+-})^{-1}$ and the first order FBC at $x = 0$ reads

$$u_t^+ + X_1^{+-}(S^{+-})^{-1}S^{++}u^+ + X_1^{+-}u^- = 0, \tag{2.21a}$$

which is different from the first order ABC (2.9a) only by the middle term.
Similarly if $(S^{-+})^{-1}$ exists, then $V^- = X_1^{-+}(S^{-+})^{-1}$ and the first order FBC at $x = 1$ is

$$u_t^- + X_1^{-+}u^+ + X_1^{-+}(S^{-+})^{-1}S^{--}u^- = 0. \tag{2.21b}$$

We shall denote these FBCs as

$$\left[(\frac{\partial}{\partial t} + \hat{X}_1)u\right]^+ = 0, \qquad x = 0, \tag{2.22a}$$
$$\left[(\frac{\partial}{\partial t} + \hat{X}_1)u\right]^- = 0, \qquad x = 1, \tag{2.22b}$$

where

$$\hat{X}_1 = \begin{pmatrix} X_1^{+-}(S^{+-})^{-1}S^{++} & X_1^{+-} \\ X_1^{-+} & X_1^{-+}(S^{-+})^{-1}S^{--} \end{pmatrix}.$$

For the second order case, we write (2.18a) in terms of the characteristic variables:

$$\begin{aligned}
&\left(I^+ + \frac{1}{s}X_1^{++} + \frac{1}{s^2}V^+S^{++} \quad , \quad \frac{1}{s}X_1^{+-} + \frac{1}{s^2}V^+S^{+-}\right)\tilde{u} \\
&= \left(I^+ + \frac{1}{s}X_1^{++} + \frac{1}{s^2}V^+S^{++} \quad , \quad \frac{1}{s}X_1^{+-} + \frac{1}{s^2}V^+S^{+-}\right) \\
&\begin{pmatrix} I^+ - \frac{1}{s}X_1^{++} - \frac{1}{s^2}[X_2 - X_1^2]^{++} & -\frac{1}{s}X_1^{+-} - \frac{1}{s^2}[X_2 - X_1^2]^{+-} \\ -\frac{1}{s}X_1^{-+} - \frac{1}{s^2}[X_2 - X_1^2]^{-+} & I^- - \frac{1}{s}X_1^{--} - \frac{1}{s^2}[X_2 - X_1^2]^{--} \end{pmatrix}\tilde{v} + O\left(\frac{1}{|s|^3}\right) \\
&= \left[I^+ + \frac{1}{s^2}(V^+S^{++} - X_2^{++})\right]\tilde{v}^+ + \frac{1}{s^2}\left[V^+S^{+-} - X_2^{+-}\right]\tilde{v}^- + O\left(\frac{1}{|s|^3}\right).
\end{aligned}$$

The optimal choice of V^+ is to annihilate the coefficient of the outgoing mode up to order $O(|s|^{-2})$. Similar computations at the right boundary condition give the analogous equation for V^-. Finally V^+ and V^- are chosen as solutions of

$$V^+S^{+-} = X_2^{+-}, \tag{2.23a}$$

$$V^-S^{-+} = X_2^{-+}. \tag{2.23b}$$

Again, if $(S^{+-})^{-1}$ and $(S^{-+})^{-1}$ exist, then the second order FBCs can be written as

$$\left[(\frac{\partial^2}{\partial t^2} + X_1\frac{\partial}{\partial t} + \hat{X}_2)u\right]^+ = 0, \qquad x = 0, \tag{2.24a}$$

$$\left[(\frac{\partial^2}{\partial t^2} + X_1\frac{\partial}{\partial t} + \hat{X}_2)u\right]^- = 0, \qquad x = 1, \tag{2.24b}$$

where

$$\hat{X}_2 = \begin{pmatrix} X_2^{+-}(S^{+-})^{-1}S^{++} & X_2^{+-} \\ X_2^{-+} & X_2^{-+}(S^{-+})^{-1}S^{--} \end{pmatrix}.$$

In the case S^{+-} and/or S^{-+} are not invertible, generalized solutions of (2.20),(2.23) have to be considered.

General cases: Considering the systems (2.20), let V^{*+} and V^{*-} denote generalized solutions of (2.20a) and (2.20b) respectively. Then we have two cases:

- $m \geq n - m$, then equation (2.20a) can be written as

$$(S^{+-})^T (V^+)^T = (X_1^{+-})^T.$$

Let $v_1^{(i)}$ and $b_1^{(i)}$ be the i^{th} columns of $(V^+)^T$ and $(X_1^{+-})^T$, respectively. Then this system is equivalent to the m underdetermined systems

$$(S^{+-})^T v_1^{(i)} = b_1^{(i)}, \quad i = 1, ..., m.$$

the solution $(v_1^*)^{(i)} \in I\!R^m$ (in the least-squares sense, that is minimizing the Euclidean norm of residuals $\left\| (S^{+-})^T v_1^{(i)} - b_1^{(i)} \right\|^2$, $\quad i = 1, ..., m$) is given by

$$(v_1^*)^{(i)} = S^{+-}((S^{+-})^T S^{+-})^{-1} b_1^{(i)}, \quad i = 1, ..., m.$$

The solution exists and is unique if S^{+-} has full rank. If S^{+-} does not have full rank, then the solution is not unique, since in this case if $(v_1^*)^{(i)}$ is a solution then the vector $(v_1^*)^{(i)} + z$ with $z \in Ker(S^{+-})$, is a solution too. A further constraint is introduced to enforce uniqueness of the solution. Typically, one requires that $(v_1^*)^{(i)}$ has minimal Euclidean norm.
On the other side, Equation (2.20b) is equivalent to $n - m$ overdetermined systems

$$(S^{-+})^T v_2^{(i)} = b_2^{(i)}, \quad i = 1, ..., n - m,$$

where $v_2^{(i)}$ and $b_2^{(i)}$ are the i^{th} columns of $(V^-)^T$ and $(X_1^{-+})^T$ respectively. The general solution is given by

$$(v_2^*)^{(i)} = (S^{-+}(S^{-+})^T)^{-1} S^{-+} b_2^{(i)}, \quad i = 1, ..., n - m.$$

- The case $m < n - m$, is similar but with $n - m$ underdetermined and m overdetermined systems.

This generalization applies to the case of (2.23).

2.2 Well-posedness of one-dimensional problem

In this section we discuss the well-posedness of the IVP

$$u_t + \Lambda(x,t) u_x + C(x,t) u = f(x,t), \qquad 0 < x < 1, \quad t \geq 0, \tag{2.25a}$$

$$u(x,0) = u_0(x), \qquad 0 < x < 1, \tag{2.25b}$$

together with boundary conditions of the form

$$\left[(\frac{\partial}{\partial t} + B(t))u\right]^{+} = 0, \qquad x = 0, \tag{2.26a}$$

$$\left[(\frac{\partial}{\partial t} + B(t))u\right]^{-} = 0, \qquad x = 1. \tag{2.26b}$$

$B(t)$ is partitioned in the same way as in Notation 2.1:

$$B(t) = \begin{pmatrix} S_0(t) & K_0(t) \\ \\ S_1(t) & K_1(t) \end{pmatrix}.$$

We assume that S_0, S_1, K_0, and K_1 are uniformly bounded for all $t \geq 0$. Clearly the first order ABCs (2.9), and the first order FBCs (2.22) are special cases of (2.26). $\Lambda(x,t), C(x,t) \in I\!R^{n\times n}$ and $f(x,t), u_0(x) \in I\!R^n$ are assumed to be C^∞-smooth functions with respect to all variables. Moreover, $f(x,t), u_0(x)$ are assumed to vanish in a neighborhood of the corners $(x,t) = (0,0)$, $(x,t) = (1,0)$.
Using this compatability assumption, we write boundary conditions (2.26) in the integral form:

$$u^+(0,t) = -\int_0^t S_0(\tau)u^+(0,\tau)d\tau - \int_0^t K_0(\tau)u^-(0,\tau)d\tau, \tag{2.27a}$$

$$u^-(1,t) = -\int_0^t S_1(\tau)u^-(1,\tau)d\tau - \int_0^t K_1(\tau)u^+(1,\tau)d\tau. \tag{2.27b}$$

Roughly speaking, the initial value problem (2.25) with boundary conditions (2.27) is called well-posed if for all smooth compatible data u_0 and f there is a unique smooth solution u, and in every finite interval $0 \leq t \leq T$ the solution can be estimated in terms of the data.
Define the outflow and inflow norms respectively by

$$\|u(t)\|_+^2 := \sum_{\lambda_j(1,t)>0} \lambda_j(1,t)|u_j(1,t)|^2 - \sum_{\lambda_j(0,t)<0} \lambda_j(0,t)|u_j(0,t)|^2,$$

and

$$\|u(t)\|_-^2 := \sum_{\lambda_j(0,t)>0} \lambda_j(0,t)|u_j(0,t)|^2 - \sum_{\lambda_j(1,t)<0} \lambda_j(1,t)|u_j(1,t)|^2.$$

Lemma 2.6. *Assume that the boundary is not characteristic and that* $\Lambda(x,t)$, $\Lambda_x(x,t), C(x,t), B(t)$ *are uniformly bounded for all* $0 \leq x \leq 1$ *and* $0 \leq t \leq T$. *Then,*

for every finite time interval $0 \le t \le T$ *there is a constant* C_T *such that, if* u *solves the IBVP (2.25),(2.27) for* $0 \le t \le T$*, then*

$$\|u(.,t)\|_2^2 + \int_0^t \left(\|u(\tau)\|_-^2 + \|u(\tau)\|_+^2\right) d\tau \ \le C_T \left(\|u_0\|_2^2 + \int_0^t \|f(.,\tau)\|_2^2 d\tau\right),$$

C_T *is independent of* f *and* u_0.

Proof.

$$\begin{aligned} \frac{d}{dt}\|u(.,t)\|_2^2 &= (u,u_t) + (u_t,u) \\ &\le -(u,\Lambda u_x) - (\Lambda u_x,u) + c_1\left\{\|u(.,t)\|_2^2 + \|f(.,t)\|_2^2\right\}. \end{aligned}$$

Integration by parts gives

$$(u,\Lambda u_x) + (u_x,\Lambda u) = \langle u,\Lambda u\rangle\,|_0^1 - (u,\Lambda_x u),$$

moreover,

$$\langle u,\Lambda u\rangle\,|_0^1 = \|u(t)\|_+^2 - \|u(t)\|_-^2.$$

Since Λ_x is uniformly bounded, we have

$$\frac{d}{dt}\|u(.,t)\|_2^2 \le \|u(t)\|_-^2 - \|u(t)\|_+^2 + c_2\left\{\|u(.,t)\|_2^2 + \|f(.,t)\|_2^2\right\}. \tag{2.28}$$

Choose

$$\rho_1 := \max_{j=1,\dots,n}\left(|\lambda_j(1,t)|, |\lambda_j(0,t)|\right), \qquad 0 \le t \le T,$$

then

$$\|u(t)\|_-^2 \le \rho_1\left(|u^+(0,t)|^2 + |u^-(1,t)|^2\right). \tag{2.29}$$

The Cauchy-Schwarz inequality for the boundary condition (2.27a) yields

$$\begin{aligned} |u^+(0,t)|^2 &\le 2\left|\int_0^t S_0(\tau)u^+(0,\tau)d\tau\right|^2 + 2\left|\int_0^t K_0(\tau)u^-(0,\tau)d\tau\right|^2 \\ &\le 2t\left(s_0^2\int_0^t \left|u^+(0,\tau)\right|^2 d\tau + k_0^2\int_0^t \left|u^-(0,\tau)\right|^2 d\tau\right), \end{aligned}$$

where $|S_0(t)| \le s_0$ and $|K_0(t)| \le k_0$ for $0 \le t \le T$.
In a similar way, the boundary condition (2.27b) can be estimated as

$$\left|u^-(1,t)\right|^2 \le 2t\left(s_1^2\int_0^t \left|u^-(1,\tau)\right|^2 d\tau + k_1^2\int_0^t \left|u^+(1,\tau)\right|^2 d\tau\right).$$

With $\hat{k} = \max(k_1, k_0)$ and $\hat{s} = \max(s_0, s_1)$, (2.29) becomes

$$\begin{aligned}\|u(t)\|_-^2 \leq\ & 2\rho_1 t \left(\hat{s}^2 \int_0^t \left(\left|u^+(0,\tau)\right|^2 + \left|u^-(1,\tau)\right|^2 \right) d\tau \right. \\ & \left. + \hat{k}^2 \int_0^t \left(\left|u^-(0,\tau)\right|^2 + \left|u^+(1,\tau)\right|^2 \right) d\tau \right). \end{aligned} \tag{2.30}$$

Consider

$$z(t) := \|u(t)\|_2^2 + \int_0^t \left(\|u(\tau)\|_-^2 + \|u(\tau)\|_+^2 \right) d\tau.$$

Using (2.28), (2.30) yields

$$\begin{aligned} z'(t) &\leq 2\|u(t)\|_-^2 + c_2 \left\{ \|u(.,t)\|_2^2 + \|f(.,t)\|_2^2 \right\} \\ &\leq 4\rho_1\rho_2 t \left(\hat{s}^2 \int_0^t \|u(\tau)\|_-^2 d\tau + \hat{k}^2 \int_0^t \|u(\tau)\|_+^2 d\tau \right) + c_2 \left\{ \|u(.,t)\|_2^2 + \|f(.,t)\|_2^2 \right\} \\ &\leq \alpha \left(\|u(.,t)\|_2^2 + \int_0^t \|u(\tau)\|_-^2 + \|u(\tau)\|_+^2 d\tau \right) + c_2 \|f(.,t)\|_2^2, \end{aligned}$$

where $\alpha := 4\rho_1\rho_2 T(s^2 + k^2) + c_2 > 0$, and

$$\rho_2 := \max_{j=1,\ldots,n} \left(|\lambda_j^{-1}(1,t)|, |\lambda_j^{-1}(0,t)| \right), \quad 0 \leq t \leq T.$$

ρ_2 is used for the estimates:
$|u^+(1,t)|^2 + |u^-(0,t)|^2 \leq \rho_2 \|u(t)\|_+^2$ and $|u^-(0,t)|^2 + |u^+(1,t)|^2 \leq \rho_2 \|u(t)\|_-^2$.
The Gronwall inequality gives the result

$$z(t) \leq C_T \left(\|u_0\|_2^2 + \int_0^t \|f(.,\tau)\|_2^2 d\tau \right),$$

where $C_T := (c_2 + 1)e^{\alpha T}$. □

Remark 2.7. *The last proof as well as the proof of Theorem 1.6 can be done under weaker regularity assumptions, namely* $f \in L^2((0,T), L^2((0,1); \mathbb{R}^n))$ *and* $u_0 \in L^2((0,1); \mathbb{R}^n)$. *First derived for the classical solution, the result for mild solutions then follows from a density argument.*

Theorem 2.8. *Under the assumptions of Lemma 2.6 and* $u_0 \in L^2((0,1); \mathbb{R}^n)$ *and* $f \in L^2((0,T), L^2((0,1); \mathbb{R}^n))$ *the IBVP (2.25),(2.27) has a unique mild solution in* $L^2((0,T), L^2((0,1); \mathbb{R}^n))$.

Proof. A fixed point method will be used to show the existence and uniqueness of the solution. For $g = (g^+, g^-) \in L^2((0,T); I\!R^n)$ solve the equation:

$$\begin{aligned} y_t + \Lambda(x,t)y_x + C(x,t)y &= f(x,t), \qquad 0 < x < 1, \;\; 0 \leq \; t \leq T, && \text{(2.31a)} \\ y(x,0) &= u_0(x), \qquad 0 \leq \; x \leq 1, && \text{(2.31b)} \\ y^+(0,t) &= g^+(t), && \text{(2.31c)} \\ y^-(1,t) &= g^-(t). && \text{(2.31d)} \end{aligned}$$

Theorem 1.6 and Remark 2.7 guarantee the existence of unique solution
$y \in C([0,T], L^2((0,1); I\!R^n))$, and $y(0,.), y(1,.) \in L^2((0,T); I\!R^n)$.
Define $Fg = ((Fg)^+, (Fg)^-)$ as

$$\begin{aligned} (Fg)^+(t) &:= -\int_0^t S_0(\tau)y^+(0,\tau)d\tau - \int_0^t K_0(\tau)y^-(0,\tau)d\tau, \qquad t \geq 0, \\ (Fg)^-(t) &:= -\int_0^t S_1(\tau)y^-(1,\tau)d\tau - \int_0^t K_1(\tau)y^+(1,\tau)d\tau, \qquad t \geq 0. \end{aligned}$$

The first to show is that F maps $L^2((0,T); I\!R^n)$ into itself:
The solution of (2.31) can be estimated, according to Theorem 1.6, for $0 \leq t \leq T$ as

$$\begin{aligned} &\|y(.,t)\|_2^2 + \int_0^t \left(|y(0,\tau)|^2 + |y(1,\tau)|^2\right) d\tau \\ &\qquad \leq K_T \left[\|u_0\|_2^2 + \int_0^t \left(\|f(.,\tau)\|_2^2 \; + |g(\tau)|^2\right) d\tau\right]. \end{aligned} \tag{2.32}$$

A similar computation as in the last proof shows that with $\alpha_1 = k_0^2 + s_0^2 + s_1^2 + k_1^2$

$$|Fg(t)|^2 \leq 2\alpha_1 t \int_0^t \left(|y(0,\tau)|^2 + |y(1,\tau)|^2\right) d\tau.$$

Integration by parts, using (2.32) and that $g \in L^2((0,T); I\!R^n)$, gives

$$\begin{aligned} \|Fg\|_2^2 &\leq \alpha_1 \left(T^2 \int_0^T \left(|y(0,t)|^2 + |y(1,t)|^2\right) dt - t^2 \left(|y(0,t)|^2 + |y(1,t)|^2\right)\right) \\ &\leq \textit{Constant}. \end{aligned}$$

The second to show is that F is contractive at least on a subinterval $(0,T_1)$ of $(0,T)$:
Given two inflow data: $g_1, g_2 \in L^2((0,T); I\!R^n)$, the difference between the corresponding outflow data can be estimated by (2.32), for $0 \leq t \leq T$ as

$$\begin{aligned} &\|(y_1 - y_2)(.,t)\|_2^2 + \int_0^t \left(|(y_1 - y_2)(0,\tau)|^2 + |(y_1 - y_2)(1,\tau)|^2\right) d\tau \\ &\qquad \leq \; K_T \int_0^t |(g_1 - g_2)(\tau)|^2 d\tau \end{aligned} \tag{2.33}$$

Now,

$$|Fg_1^+(t) - Fg_2^+(t)|^2 \leq 2t(s_0^2 + k_0^2) \int_0^t |(y_1 - y_2)(0,\tau)|^2 d\tau,$$

$$|Fg_1^-(t) - Fg_2^-(t)|^2 \leq 2t(s_1^2 + k_1^2) \int_0^t |(y_1 - y_2)(1,\tau)|^2 d\tau.$$

Using (2.33) we get

$$|Fg_1(t) - Fg_2(t)|^2 \leq 2t\alpha_1 \int_0^t \left(|(y_1 - y_2)(0,\tau)|^2 + |(y_1 - y_2)(1,\tau)|^2\right) d\tau.$$

Integration by parts gives

$$\begin{aligned}\|Fg_1 - Fg_2\|_2^2 \leq \alpha_1 \{T^2 &\int_0^T \left(|(y_1 - y_2)(0,t)|^2 + |(y_1 - y_2)(1,t)|^2\right) dt \\ &-t^2 \left(|(y_1 - y_2)(0,t)|^2 + |(y_1 - y_2)(1,t)|^2\right)\} \\ &\leq \alpha_1 K_T T^2 \|g_1 - g_2\|_2^2.\end{aligned}$$

F is contraction for $T_1 < \dfrac{1}{\sqrt{\alpha_1 K_T}}$.
The contractivity of F depends only on $\alpha_1 K_T$ (but not on the initial condition or the inhomogeneity term). So we can apply the iteration first on a subinterval $(0, T_1)$ of $(0, T)$, then $(T_1, 2T_1)$ and so on. The local solution can be continued in t to reach T. □

This theorem shows that our problem is strongly well-posed in the sense of Kreiss [**47**].

2.3 Convergence to the steady state

We now consider the convergence of the IVP (2.1)-(2.2) with first order FBCs (2.17) as $t \to \infty$ to the solution of the corresponding steady problem. The non-singular matrix V in (2.17) is used to accelerate the convergence to the steady state and it does not effect the convergence itself. For convenience we take V^+ and V^- as identity matrices. Then, we have

$$u_t + \Lambda u_x + Cu = f(x), \qquad 0 < x < 1, \quad t > 0, \tag{2.34a}$$

$$u(x,0) = u_0(x), \qquad 0 < x < 1, \tag{2.34b}$$

$$\left[(\frac{\partial}{\partial t} + S)u\right]^+ = 0, \qquad x = 0, \tag{2.34c}$$

$$\left[(\frac{\partial}{\partial t} + S)u\right]^- = 0, \qquad x = 1. \tag{2.34d}$$

The corresponding steady state problem reads

$$\Lambda u_x^* + Cu^* = f(x), \qquad 0 < x < 1, \tag{2.35a}$$

with transparent boundary conditions

$$(Su^*)^+ = 0, \qquad x = 0, \tag{2.35b}$$
$$(Su^*)^- = 0, \qquad x = 1. \tag{2.35c}$$

Define:

$$q(x,t) := u(x,t) - u^*(x),$$

then q satisfies

$$\begin{aligned} q_t + \Lambda q_x + Cq &= 0, && 0 < x < 1,\ t \geq 0, && (2.36a)\\ q(x,0) &= u_0(x) - u^*(x) =: q_0(x), && 0 < x < 1, && (2.36b)\\ \left[(\frac{\partial}{\partial t} + S)q\right]^+ &= 0, && x = 0, && (2.36c)\\ \left[(\frac{\partial}{\partial t} + S)q\right]^- &= 0, && x = 1. && (2.36d) \end{aligned}$$

By taking the Laplace transform (with dual variable s) of (2.36) we obtain the ordinary differential equation

$$\begin{aligned} s\tilde{q} + \Lambda \tilde{q}_x + C\tilde{q} &= q_0(x), && 0 < x < 1, && (2.37a)\\ s\tilde{q}^+ + (S\tilde{q})^+ &= q_0^+(0), && x = 0, && (2.37b)\\ s\tilde{q}^- + (S\tilde{q})^- &= q_0^-(1), && x = 1. && (2.37c) \end{aligned}$$

The following lemma, which relates the asymptotic behavior of the original function with the limit of the image function, will be used to prove the main theorem of this section.

Lemma 2.9. **[21]** *Suppose that $b(t)$ belongs to a Banach space with norm $\|.\|$ and $\tilde{b}(s)$ is its Laplace transform. Then we have*

$$\lim_{t\to\infty} b(t) = \lim_{s\to 0^+} s\tilde{b}(s),$$

provided that $\lim_{t\to\infty} b(t)$ exists.

Proof. Let $\lim_{t\to\infty} b(t) = a$, then for any fixed $\epsilon > 0$, there exists a T, such that $\|b(t) - a\| < \epsilon \quad$ for $t > T$.
Define $r(t) := b(t) - a$. For all $s > 0$ consider

$$\begin{aligned}
\|\tilde{b}(s) - \frac{a}{s}\| &= \|\int_0^\infty e^{-st} b(t)dt - \frac{a}{s}\| \\
&= \|\int_0^\infty e^{-st} r(t)dt\| \\
&\leq \int_0^T e^{-st}\|r(t)\|dt + \int_T^\infty e^{-st}\|r(t)\|dt \\
&< \int_0^T \|r(t)\|dt + \epsilon\int_T^\infty e^{-st}dt \\
&\leq c_1 + \frac{\epsilon}{s}e^{-sT} \leq c_1 + \frac{\epsilon}{s}
\end{aligned}$$

$$\Rightarrow \|s\tilde{b}(s) - a\| < \epsilon + sc_1, \qquad \forall s > 0, \quad c_1 \in I\!R^+,$$

but with

$$s \in (0, \frac{\epsilon}{c_1}],$$

we have

$$\|s\tilde{b}(s) - a\| < 2\epsilon$$

for $s \to 0^+$ gives the result. □

Theorem 2.10. *If the solution $q(x,t)$ of (2.36) converges in $L^2(0,1)$ as $t \to \infty$, then it converges to zero in $L^2(0,1)$.*

Proof. To make use of the previous Lemma we need to show that $\tilde{q}(x,s)$ defined in (2.37) is a natural extension of the Laplace transform near $s = 0$. That is, $\tilde{q}(x,s)$ is an analytic function of s in the neighborhood of 0.
Suppose that

$$Z := H^1((0,1); I\!R^n) \times I\!R^n, \qquad Y := L^2((0,1); I\!R^n) \times I\!R^n.$$

Define the operator L as

$$L : Y \supset Z \supset H^1(0,1)^n =: D(L) \ni \tilde{q} \longmapsto \begin{pmatrix} (\Lambda\partial x + C)\tilde{q} \\ (S\tilde{q})^+(0) \\ (S\tilde{q})^-(1) \end{pmatrix} \in Y.$$

Accordingly, (2.37) can be written as

$$(L+sI)\begin{pmatrix}\tilde{q}\\ \tilde{q}^+(0)\\ \tilde{q}^-(1)\end{pmatrix}=\begin{pmatrix}q_0(x)\\ q_0^+(0)\\ q_0^-(1)\end{pmatrix},\qquad \tilde{q}^+(0)\in I\!R^m,\qquad \tilde{q}^-(1)\in I\!R^{n-m}. \quad (2.38)$$

Focusing on the operator L, we notice that it has some useful properties:
First, L is invertible. Consider

$$q_1=\begin{pmatrix}f(x)\\ \alpha^+\\ \alpha^-\end{pmatrix}\in Y,\qquad \alpha^+\in I\!R^m,\quad \alpha^-\in I\!R^{n-m},$$

and search for $\tilde{q}\in D(L)$, such that

$$\begin{aligned}\Lambda\tilde{q}_x+C\tilde{q} &= f(x), \qquad 0<x<1,\\ (S\tilde{q})^+(0) &= \alpha^+,\\ (S\tilde{q})^-(1) &= \alpha^-.\end{aligned}$$

This inhomogeneous boundary value problem is equivalent to

$$\begin{aligned}&\breve{w}_x+R\breve{w}=h(x), \qquad 0<x<1,\\ &\breve{w}^+(0)=0,\qquad \breve{w}^-(1)=0,\end{aligned} \quad (2.40)$$

where $\breve{w}=S\tilde{q}-\begin{pmatrix}\alpha^+\\ \alpha^-\end{pmatrix}$, $h(x)=S\Lambda^{-1}f(x)-(I+R)\begin{pmatrix}\alpha^+\\ \alpha^-\end{pmatrix}$, and S is defined as in (2.14). The last equation (2.40) has the unique solution

$$\breve{w}^+(x)=\int_0^x e^{-R^+(x-y)}h^+(y)dy, \qquad 0<x<1,$$

$$\breve{w}^-(x)=-\int_x^1 e^{-R^-(y-x)}h^-(y)dy, \qquad 0<x<1.$$

Second, L^{-1} is a bounded operator from Y to $D(L)$. Multiplying (2.37a) by $\tilde{q}$ and integrating

$$(\Lambda\tilde{q}_x,\tilde{q})+(\tilde{q},\Lambda\tilde{q}_x)+(C\tilde{q},\tilde{q})+(\tilde{q},C\tilde{q})=2(f,\tilde{q}). \quad (2.41)$$

Integration by parts gives

$$(\tilde{q},\Lambda\tilde{q}_x)+(\tilde{q}_x,\Lambda\tilde{q})=\langle\tilde{q},\Lambda\tilde{q}\rangle\,\big|_0^1,$$

using the properties of C

$$(C\tilde{q},\tilde{q})+(\tilde{q},C\tilde{q})=(\tilde{q},(C+C^T)\tilde{q})\geq\ 2\delta\|\tilde{q}\|_2^2>0,\quad \delta>0,$$

and that

$$2(f,\tilde{q})\leq\frac{2}{\delta}\|f\|_2^2+\frac{\delta}{2}\|\tilde{q}\|_2^2.$$

Thus, (2.41) gives

$$\langle\tilde{q},\Lambda\tilde{q}\rangle\,|_0^1+\frac{3\delta}{2}\|\tilde{q}\|_2^2\leq\frac{2}{\delta}\|f\|_2^2. \tag{2.42}$$

Choose

$$\lambda_M:=\max_{j=1,\ldots,n}\{|\lambda_j|\},$$

then

$$\begin{aligned}\langle\tilde{q},\Lambda\tilde{q}\rangle\,|_0^1 &= \sum_{j=1}^{m}\lambda_j(|\tilde{q}_j(1)|^2-|\tilde{q}_j(0)|^2)+\sum_{j=m+1}^{n}\lambda_j(|\tilde{q}_j(1)|^2-|\tilde{q}_j(0)|^2)\\ &\geq -\sum_{j=1}^{m}\lambda_j|\tilde{q}_j(0)|^2+\sum_{j=m+1}^{n}\lambda_j|\tilde{q}_j(1)|^2\\ &\geq -n\lambda_M(|\alpha^+|^2+|\alpha^-|^2).\end{aligned}$$

Substituting the result in (2.42), we get

$$\|\tilde{q}\|_2^2\leq C(\|f\|_2^2+|\alpha^+|^2+|\alpha^-|^2). \tag{2.43}$$

But

$$\tilde{q}_x=-\Lambda^{-1}C\tilde{q}+\Lambda^{-1}f,$$

then

$$\|\tilde{q}_x\|_2^2\leq C_1(\|\tilde{q}\|_2^2+\|f\|_2^2). \tag{2.44}$$

Using (2.43) and (2.44) we get

$$\|\tilde{q}\|_{D(L)}^2\leq C(\|f\|_2^2+|\alpha^+|^2+|\alpha^-|^2)=C\|q_1\|_Y. \tag{2.45}$$

Now, define $L_p^{-1}:=P\circ L^{-1}:Y\longmapsto Y$, where $P:Z\longmapsto Y$ is the identity compact operator. Since $L^{-1}:Y\longmapsto D(L)$ is bounded, then L_p^{-1} is compact.
Choose s small enough such that

$$|s|<\|L_p^{-1}\|^{-1}, \tag{2.46}$$

then $(I + sL_p^{-1})$ is an invertible operator from Y to Y. Moreover, the resolvent function $(I + sL_p^{-1})^{-1}$ is an analytic function of s with norm that satisfies

$$\|(\frac{I}{s} + L_p^{-1})^{-1}\|_Y \leq \frac{1}{\frac{1}{s} - \|L_p^{-1}\|_Y},$$

and so in particular

$$\|(\frac{I}{s} + L_p^{-1})^{-1}\|_Y \to 0 \quad as\ s \to 0. \tag{2.47}$$

Consider the first (main) part of (2.38)

$$(L + sI)\tilde{q} = q_0, \tag{2.48}$$

and let s be as in (2.46). Take L_p^{-1} of both sides of (2.48), then

$$(I + sL_p^{-1})\tilde{q} = L_p^{-1}q_0.$$

Since $(I + sL_p^{-1})$ is invertible, we obtain

$$\tilde{q} = (I + sL_p^{-1})^{-1}L_p^{-1}q_0,$$

which is analytic in s and bounded. The multiplication of the last equation by s yields

$$s\tilde{q} = (\frac{I}{s} + L_p^{-1})^{-1}L_p^{-1}q_0.$$

Using (2.47) and that L_p^{-1} is bounded, gives that $\|s\tilde{q}(x,s)\|_2 \to 0 \ \ as \ \ s \to 0$.
From Lemma 2.9 we obtain the result. □

Chapter 3

Numerical Approximation

Consider the first order hyperbolic system in characteristic form

$$u_t + \Lambda u_x + Cu = f(x,t), \tag{3.1a}$$

in the stripe $0 < x < 1,\ t > 0$. Here, Λ and C are constant $n \times n$ matrices and Λ is partitioned as in (2.3). The solution of (3.1a) is uniquely determined if we prescribe initial values for $t = 0$:

$$u(x,0) = u_0(x), \qquad 0 \le x \le 1, \tag{3.1b}$$

and boundary conditions at $x = 0, 1$:

$$u_t^+ + (Su)^+ = 0, \qquad x = 0, \tag{3.1c}$$

$$u_t^- + (Su)^- = 0, \qquad x = 1. \tag{3.1d}$$

Here, S is defined as in (2.14). While (3.1c)-(3.1d) represent general FBCs. Furthermore, the support of f and u_0 are assumed to be in $(0,1)$.
We want to solve the above problem by finite difference approximation. For that reason, we introduce a mesh size $h := \triangle x$, a time step $k := \triangle t$, and discretize the $(x,t)-$stripe $[0,1] \times I\!R_0^+$ using the mesh points

$$x_j = jh, \qquad j = 0, 1, 2, ..., J$$

$$t_l = lk, \qquad l = 0, 1, 2, ...$$

We assume that $r := k/h$ is constant and use the notation $u_j^l \in I\!R^n$ to approximate the exact solution u at (x_j, t_l). According to the partition of u, we set

$$u_j^l = \begin{pmatrix} (u^+)_j^l \\ \\ (u^-)_j^l \end{pmatrix}, \quad \left(u^+\right)_j^l \in I\!R^m, \quad \left(u^-\right)_j^l \in I\!R^{n-m}.$$

The outline of this chapter is as follows: In Section 3.1, we introduce a finite difference scheme to solve the IBVP (3.1). In Section 3.2, we apply the well known stability theory due to Gustafsson, Kreiss, and Sundström (GKS-theory) to prove the stability of this scheme. Two numerical examples are given in Section 3.3, in the first example, we discuss briefly a 2×2 model system and prove the convergence of this system with first order FBCs to the correct steady state. In the second example, we compare the numerical approximations for different choices of the scaling matrices for a 3×3 system.

3.1 Numerical scheme

Lax-Wendorf scheme (LW-scheme) based on the expansion

$$u(x, t+k) = u(x,t) + k u_t(x,t) + \frac{k^2}{2} u_{tt}(x,t) + O(k^3), \tag{3.2}$$

where u_{tt} can be determined using (3.1a) as follows

$$\begin{aligned} u_{tt} &= (-\Lambda u_x - Cu + f)_t \\ &= -\Lambda u_{tx} - C u_t + f_t \\ &= \Lambda(\Lambda u_x + Cu - f)_x + C(\Lambda u_x + Cu - f) + f_t \\ &= \Lambda^2 u_{xx} + (\Lambda C + C\Lambda) u_x + C^2 u + f_t - \Lambda f_x - Cf. \end{aligned}$$

Substituting the above equation into (3.2) yields

$$\begin{aligned} u(x, t+k) &= u(x,t) - k(\Lambda u_x(x,t) + Cu(x,t) - f(x,t)) \\ &\quad + \frac{k^2}{2}(\Lambda^2 u_{xx}(x,t) + (\Lambda C + C\Lambda) u_x(x,t) + C^2 u(x,t) \\ &\quad + f_t(x,t) - \Lambda f_x(x,t) - Cf(x,t)) + O(k^3). \end{aligned}$$

The LW-scheme uses centered differences to approximate the spatial derivatives of u. Furthermore, the derivatives of f will be appropriately discertized. The resulting

scheme will be then

$$\begin{aligned} u_j^{l+1} &= u_j^l - \frac{1}{2}r\Lambda(u_{j+1}^l - u_{j-1}^l) - kCu_j^l + \frac{1}{2}(r\Lambda)^2(u_{j+1}^l - 2u_j^l + u_{j-1}^l) \\ &+\frac{1}{4}rk(\Lambda C + C\Lambda)(u_{j+1}^l - u_{j-1}^l) + \frac{1}{2}(kC)^2u_j^l + \frac{1}{2}k(f_j^{l+1} + f_j^l) \\ &-\frac{1}{4}rk\Lambda(f_{j+1}^l - f_{j-1}^l) - \frac{1}{2}k^2Cf_j^l, \qquad l = 0, 1, ..., \quad j = 1, ..., J-1. \end{aligned} \tag{3.3a}$$

To solve (3.3a) uniquely, we provide initial values

$$u_j^0 = u_0(x_j), \quad j = 0, 1, 2, ..., J, \tag{3.3b}$$

and specify at each time level $t_l = lk,\ l = 1, 2, ...$, boundary values u_0^{l+1}, u_J^{l+1}. These boundary values split into two groups: The first group, which we refer to as inflow boundary conditions is

$$\left(u^+\right)_0^{l+1}, \quad \left(u^-\right)_J^{l+1}.$$

The second group is

$$\left(u^-\right)_0^{l+1}, \quad \left(u^+\right)_J^{l+1},$$

which we refer to as the outflow boundary conditions.
The inflow values are determined by the discretization of the boundary conditions (3.1c)-(3.1d), while the outflow values are obtained by introducing numerical boundary conditions. In this work we shall consider two types of numerical boundary conditions, the first type is upwinding in which u at the boundaries satisfy the homogeneous version of the system (3.1a), and the second type is first order extrapolation.

Definition 3.1. *The general horizontal extrapolation of order q for the outflow data u^- at $x = 0$ is*

$$(E_+ - I)^{q+1}\left(u^-\right)_0^{l+1} = 0, \quad q = 0, 1, ...,$$

and that of u^+ at $x = 1$

$$\left(I - E_+^{-1}\right)^{q+1}\left(u^+\right)_J^{l+1} = 0, \quad q = 0, 1, ...,$$

where $E_+u_j := u_{j+1}$.

Using boundary condition (3.1c), we write

$$D_+^t\left(u^+\right)_0^l + \left((Su)^+\right)_0^l = 0,$$

which gives

$$\left(u^+\right)_0^{l+1} = \left(u^+\right)_0^l - k\left((Su)^+\right)_0^l. \tag{3.4}$$

Since f is compactly supported in $(0,1)$, the outflow part of (3.1a) at $x=0$ satisfies

$$u_t^- + \Lambda^- u_x^- + (Cu)^- = 0,$$

which is discretized as

$$D_+^t\left(u^-\right)_0^l + \Lambda^- D_+^x\left(u^-\right)_0^l + \left((Cu)^-\right)_0^l = 0.$$

Hence

$$\left(u^-\right)_0^{l+1} = \left(I + r\Lambda^-\right)\left(u^-\right)_0^l - k\left((Cu)^-\right)_0^l - r\Lambda^-\left(u^-\right)_1^l. \tag{3.5a}$$

An alternative numerical boundary condition is the first order extrapolation

$$\left(u^-\right)_0^{l+1} = 2\left(u^-\right)_1^{l+1} - \left(u^-\right)_2^{l+1}. \tag{3.5b}$$

The discretization of the right boundary conditions is treated in a similar manner.

3.2 Stability of the finite difference scheme

In solving linear hyperbolic partial differential equations numerically by means of finite difference approximations, a principal difficulty both theoretically and in practice is the question of stability. For the "Cauchy problem" on the unbounded domain $(-\infty,\infty)$, a fairly complete stability theory based on the Fourier analysis has been worked out during the last few decades by von Neumann, Lax, Kreiss, and others [**70, 72, 76**]. For the "initial boundary value problem" on a domain such as $[0,\infty)$ or $[0,1]$, however, Fourier analysis cannot be applied in a straightforward way, and progress has been slower and technically more complex. Important contributions in this area were made by S. Osher [**67**] and by H.-O. Kreis [**60**], and are based on various kinds of normal mode analysis that extend the Fourier methods. A comprehensive theory of this type was presented in an influential paper by Gustafsson, Kreiss, and Sundström (briefly: GKS) [**37**]. The complicated algebraic conditions of the GKS-theory were simplified in following work of Goldberg and Tadmor [**38**].
In this section we apply the GKS-theory to show the stability of the difference approximation (3.3)-(3.5a)(or(3.5b)) and the corresponding right boundary discretization. We intend to provide both sufficient and necessary conditions for the stability of this discrete IBVP. It appears that the IBVP does not have the standard form

presented in the GKS-theory and thus, this stability theory is not directly applicable.
The discrete IBVP under consideration is given with two boundaries. According to the Theorem 3.2 below, which is valid for any of the GKS stability definitions, it is sufficient to consider the problem on the positive plane $x \geq 0$, i.e., on the index range $j \geq 0$.

Theorem 3.2. *[**37**], Thm. 5.4] Consider the difference approximation for $t \geq 0$ and $0 \leq x \leq 1$ and assume that the corresponding left and right quarter-plane problems (which we get by removing one boundary to infinity) are stable, then the original problem is also stable.*

The idea behind the theorem is that the basic difference scheme (3.3) and each of the boundary conditions separated into the two quarter plane problems that are relatively nice to handle. Therefore, we will consider only the stability of the right quarter plane problem, while the left quarter one is analogue.
To fit our approximation into the form discussed in [**37**], we write (3.3) as

$$u_j^{l+1} = Qu_j^l + kb_j^l, \tag{3.6a}$$
$$u_j^0 = u_0(x_j), \quad j = 0, 1, 2, ..., \tag{3.6b}$$

where

$$Q = \sum_{\sigma=-1}^{1} \Lambda_\sigma E_+^\sigma, \quad E_+ u_j = u_{j+1},$$

$$\Lambda_0 = I - kC - (r\Lambda)^2 + \frac{1}{2}(kC)^2,$$
$$\Lambda_{\pm 1} = \mp\frac{1}{2}r\Lambda + \frac{1}{2}(r\Lambda)^2 \pm \frac{1}{4}rk(\Lambda C + C\Lambda),$$
$$b_j^l = \frac{1}{2}\left(f_j^{l+1} + f_j^l\right) - \frac{1}{4}r\Lambda\left(f_{j+1}^l - f_{j-1}^l\right) - \frac{1}{2}kCf_j^l.$$

The boundary values are written as

$$u_0^{l+1} = B_{0,0}u_0^l + B_{1,0}u_1^l + B_{1,1}u_1^{l+1} + B_{2,1}u_2^{l+1}, \tag{3.7}$$

where the above matrices are determined by the numerical boundary conditions under consideration. For the upwinding case (3.4)-(3.5a), we have

$$B_{0,0} = \begin{pmatrix} I^+ & 0 \\ 0 & I + r\Lambda^- \end{pmatrix} - k \begin{pmatrix} S^{++} & S^{+-} \\ C^{-+} & C^{--} \end{pmatrix},$$

$$B_{1,0} = \begin{pmatrix} 0 & 0 \\ 0 & -r\Lambda^- \end{pmatrix}, \qquad B_{1,1} = B_{2,1} = 0. \tag{3.8a}$$

However, if extrapolation (3.4)-(3.5b) is used, then

$$B_{0,0} = \begin{pmatrix} I^+ & 0 \\ 0 & 0 \end{pmatrix} - k \begin{pmatrix} S^{++} & S^{+-} \\ 0 & 0 \end{pmatrix},$$

$$B_{1,1} = \begin{pmatrix} 0 & 0 \\ 0 & -2I^- \end{pmatrix}, \quad B_{2,1} = \begin{pmatrix} 0 & 0 \\ 0 & -I^- \end{pmatrix}, \quad B_{1,0} = 0. \tag{3.8b}$$

There are different ways to define stability of finite difference schemes. GKS [**37**] discussed some possible definitions of which we choose the one that allows us to make use of the available results.

Let $l^2(x)$ denote the space of all grid functions $u_j = u(x_j),\ x_j = jh,\ j = 0, 1, ...,$ with $\sum_{j=0}^{\infty} |u_j|^2 < \infty$ and define the scalar product and norm by

$$(u, v)_h = \sum_{j=0}^{\infty} h u_j^* v_j, \qquad \|u\|_h^2 = (u, u)_h.$$

We define $l^2(t)$ and $l^2(x, t)$ in the corresponding way and denote by

$$(u, v)_k = \sum_{l=0}^{\infty} k u^*(t_l) v(t_l), \qquad \|u\|_k^2 = (u, u)_k,$$

$$(u, v)_{h,k} = \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} hk u_j^*(t_l) v_j(t_l), \qquad \|u\|_{h,k}^2 = (u, u)_{h,k},$$

the corresponding norms and scalar products.

Definition 3.3. *[*[**37**]*, Def. 3.3] Assume that the initial function is zero. The difference scheme (3.6)-(*3.8a*)(or(*3.8b*)) is stable, if there exist constants $c_0 > 0, \alpha_0 \geq 0$ such that, for all $t = t_l = lk$, all $\alpha > \alpha_0$, and all h, an estimate*

$$\left(\frac{\alpha - \alpha_0}{\alpha k + 1}\right) \left\|e^{-\alpha t} u_0\right\|_k^2 + \left(\frac{\alpha - \alpha_0}{\alpha k + 1}\right)^2 \left\|e^{-\alpha t} u\right\|_{h,k}^2 \leq c_0 \left\|e^{-\alpha(t+k)} b\right\|_{h,k}^2$$

holds.

While here the vector b_j^l of the basic scheme (3.6a) is a general combination of f and its derivatives, in [**37**] we have $b_j^l = f_j^l$. However, Goldberg *et al.* [**38**] showed that this generalization does not affect the results of [**37**] and they raised the question of stability in the sense of Definition 3.3.

The definition of stability for the difference scheme for the left quarter plane problem is the same, except that the norm is taken over the grid on $(-\infty, 1]$ and u_0 is replaced by u_J.

In the following, we shall reduce the above stability question to that of the principal part of the scalar outflow approximations, i.e., the part obtained by eliminating the terms of order k, k^2, and all inhomogeneity vectors. This result is based on Theorem 4.3 of [**37**], which provides a necessary and sufficient determinantal stability criterion given entirely in terms of the principal part of the approximations. The mere existence of such a criterion implies that for the stability purposes we may consider a basic scheme of (3.6)-(3.8a)(or(3.8b)) of the form

$$u_j^{l+1} = \tilde{Q} u_j^l, \quad \tilde{Q} = \sum_{\sigma=-1}^{1} \tilde{\Lambda}_\sigma E_+^\sigma, \quad E_+ u_j = u_{j+1}, \tag{3.9}$$

where

$$\begin{aligned} \tilde{\Lambda}_0 &= I - (r\Lambda)^2, \\ \tilde{\Lambda}_{\pm 1} &= \mp \frac{1}{2} r\Lambda + \frac{1}{2} (r\Lambda)^2, \end{aligned}$$

and the boundary conditions

$$\left(u^+\right)_0^{l+1} = \left(u^+\right)_0^l, \tag{3.10}$$

$$\left(u^-\right)_0^{l+1} = \left(I + r\Lambda^-\right) \left(u^-\right)_0^l - r\Lambda^- \left(u^-\right)_1^l, \tag{3.11a}$$

$$\left(u^-\right)_0^{l+1} = 2 \left(u^-\right)_1^{l+1} - \left(u^-\right)_2^{l+1}. \tag{3.11b}$$

The scheme (3.9) is now consistent with

$$u_t + \Lambda u_x = 0.$$

We split the basic scheme and the boundary values into inflow and outflow parts respectively

$$\begin{aligned} \left(u^-\right)_j^{l+1} &= \left(u^-\right)_j^l - \frac{r\Lambda^-}{2} \left(\left(u^-\right)_{j+1}^l - \left(u^-\right)_{j-1}^l \right) \\ &+ \frac{(r\Lambda^-)^2}{2} \left(\left(u^-\right)_{j+1}^l - 2 \left(u^-\right)_j^l + \left(u^-\right)_{j-1}^l \right), \end{aligned} \tag{3.12}$$

$$\left(u^-\right)_0^{l+1} = \left(I + r\Lambda^-\right)\left(u^-\right)_0^l - r\Lambda^-\left(u^-\right)_1^l, \tag{3.13a}$$

$$\left(u^-\right)_0^{l+1} = 2\left(u^-\right)_1^{l+1} - \left(u^-\right)_2^{l+1}, \tag{3.13b}$$

and

$$\begin{aligned}\left(u^+\right)_j^{l+1} &= \left(u^+\right)_j^l - \frac{r\Lambda^+}{2}\left(\left(u^+\right)_{j+1}^l - \left(u^+\right)_{j-1}^l\right) \\ &+ \frac{\left(r\Lambda^+\right)^2}{2}\left(\left(u^+\right)_{j+1}^l - 2\left(u^+\right)_j^l + \left(u^+\right)_{j-1}^l\right),\end{aligned} \tag{3.14}$$

$$\left(u^+\right)_0^{l+1} = \left(u^+\right)_0^l. \tag{3.15}$$

Obviously, (3.6)-(3.8a)(or(3.8b)) is stable if and only if both parts are stable. Before we proceed, we include the following assumptions that are necessary for the result contained in this section.

Assumption 3.4. *1. The associated initial value scheme is stable.*

2. The difference scheme is either dissipative or nondissipative.

A necessary condition for the stability of the initial value scheme is to satisfy the CFL(Courant-Friedrichs-Levy)-condition. CFL-condition simply asserts that the analytical domain of dependence is contained in the numerical domain of dependence. For the LW-scheme, this gives

$$\max_{\nu=1,\ldots,n} |\lambda_\nu r| \leq 1, \tag{3.16}$$

Definition 3.5. **[76]** *The difference scheme (3.9) is dissipative of order* $2s$ *if there exists* $c > 0$ *such that the eigenvalues* $\mu_\nu(\xi)$ *of the amplification matrix of* $\tilde{Q}$ *satisfies the following estimate*

$$|\mu_\nu(\theta)|^2 \leq 1 - c\,|\theta|^{2s}, \qquad |\theta| \leq \pi.$$

This condition is equivalent to (see **[72]**)

$$|\mu_\nu(\theta)|^2 \leq 1 - \acute{c}\sin^{2s}(\theta/2), \quad \acute{c} > 0.$$

The amplification matrix of $\tilde{Q}$ reads

$$I - ir\Lambda\sin\theta - (r\Lambda)^2(1-\cos\theta),$$

with eigenvalues

$$\mu_\nu(\theta) = 1 - ir\lambda_\nu \sin\theta - 2r^2\lambda_\nu^2 \sin^2(\theta/2), \qquad \nu = 1, ..., n.$$

This gives

$$\begin{aligned} |\mu_\nu(\theta)|^2 &= [1 - 2(r\lambda_\nu)^2 \sin^2(\theta/2)]^2 + (r\lambda_\nu)^2 \sin^2\theta \\ &= 1 - 2(r\lambda_\nu)^2[4\sin^2(\theta/2) + 4(r\lambda_\nu)^2 \sin^4(\theta/2) + \sin^2\theta] \\ &= 1 - 4(r\lambda_\nu)^2[1 - (r\lambda_\nu)^2] \sin^4(\theta/2), \qquad \nu = 1, ..., n. \end{aligned} \tag{3.17}$$

Thus, the difference scheme (3.9) is dissipative of order 4 if r is chosen to satisfy

$$0 < |\lambda_\nu r| \leq 1, \qquad \nu = 1, ..., n.$$

Since Λ is regular, Assumption 3.4 is fulfilled if the CFL-condition (3.16) is satisfied. We split the outflow approximation (3.12)-(3.13a)(or(3.13b) into $n - m$ scalar components, each of the form

$$\begin{aligned} v_j^{l+1} &= v_j^l - \frac{\kappa}{2}(v_{j+1}^l - v_{j-1}^l) + \frac{\kappa^2}{2}(v_{j+1}^l - 2v_j^l + v_{j-1}^l) \\ &= \frac{1}{2}(\kappa^2 + \kappa)v_{j-1}^l + (1 - \kappa^2)v_j^l + \frac{1}{2}(\kappa^2 - \kappa)v_{j+1}^l \end{aligned} \tag{3.18}$$

where $\kappa := r\lambda_\nu$, for fixed $\lambda_\nu \in \Lambda^-$, and

$$v_0^{l+1} = v_0^l - \kappa(v_1^l - v_0^l). \tag{3.19a}$$

or

$$v_0^{l+1} = 2v_1^{l+1} - 2v_2^{l+1}. \tag{3.19b}$$

The scheme (3.6)-(3.8a)(or(3.8b)) is stable if and only if (3.12)-(3.13a)(or(3.13b)) and (3.14)-(3.15) are stable, and the latter are stable if and only if their scalar components are. Lemma 2.3 of [**38**] shows that the scalar components of the inflow approximation (3.14)-(3.15) are stable (for $0 < \kappa \leq 1$). So we conclude the main result of this section

Lemma 3.6. *The Approximation (3.6)-(*3.8a*)(or(*3.8b*)) is stable if and only if the scalar outflow components (3.18)-(*3.19a*)(or(*3.19b*)) are stable.*

To discuss the stability of (3.18)-(3.19a)(or(3.19b)) we use the discrete Laplace transform, which is one of the few approaches available for analyzing the stability of difference schemes for initial boundary value problems. This approach is used to transform out the temporal differences (time derivatives) and consider the scheme in transform space as a difference scheme in j.

Definition 3.7. *The discrete Laplace transform of* $u = \{u^l\}$ *is the function* $\tilde{u} := \mathcal{L}(\{u^l\})$ *defined by*

$$\tilde{u}(z) := \sum_{l=0}^{\infty} e^{-zl} u^l$$

where $z \in \mathbb{C},\ \Re z > 0$ *and* $\Im z \in [-\pi, \pi]$.

We take the discrete Laplace transform of equation (3.18)-(3.19a)(or(3.19b)) and obtain the resolvent equation

$$z\tilde{v}_j = \frac{1}{2}(\kappa^2 + \kappa)\tilde{v}_{j-1} + (1 - \kappa^2)\tilde{v}_j + \frac{1}{2}(\kappa^2 - \kappa)\tilde{v}_{j+1}, \tag{3.20}$$

and the transformed boundary conditions

$$z\tilde{v}_0 = \tilde{v}_0 - \kappa(\tilde{v}_1 - \tilde{v}_0), \tag{3.21a}$$

$$z(\tilde{v}_0 - 2\tilde{v}_1 + \tilde{v}_2) = 0. \tag{3.21b}$$

Definition 3.8. *The complex number* z, $|z| > 1$, *is an eigenvalue of equations (3.20)-(*3.21a*)(or(* 3.21b*)) if*

1. *there exists a vector* $\tilde{v} = [\tilde{v}_0\ \tilde{v}_1\ ...]^T$ *such that* $(z, \tilde{v})$ *satisfies equations (3.20)-(*3.21a*)(or(* 3.21b*)), and*
2. $\|\tilde{v}\|_h < \infty$.

Definition 3.9. *The complex number* z *is a generalized eigenvalue of equations (3.20)-(*3.21a*)(or(* 3.21b*)) if*

1. *there exists a vector* $\tilde{v} = [\tilde{v}_0\ \tilde{v}_1\ ...]^T$ *such that* $(z, \tilde{v})$ *satisfies equations (3.20)-(*3.21a*)(or(* 3.21b*)),*
2. $|z| = 1$, *and*
3. $\tilde{v}_k$ *satisfies*

$$\tilde{v}_k(z) = \lim_{\omega \to z, |\omega| > 1} \tilde{v}_k(\omega),$$

where $(\omega, \tilde{v}(\omega))$ *is a solution to equation (3.20).*

The result from **[36]** is given in the following proposition.

Proposition 3.10. *The difference scheme (3.18)-(*3.19a*)(or(*3.19b*)) is stable if and only if the eigenvalue problem (3.20)-(*3.21a*)(or(* 3.21b*)) has no eigenvalues and no generalized eigenvalues.*

Theorem 3.11. *The approximation (3.18) in combination with one of the boundary conditions (*3.19a*) or (*3.19b*) is stable for* $-1 \leq \kappa < 0$.

To prove this theorem we apply Proposition 3.10 and the first part of the following lemma, which describes the root of the characteristic equation of (3.20)

$$zk = k + \frac{\kappa}{2}(k^2 - 1) + \frac{\kappa^2}{2}(k-1)^2. \tag{3.22}$$

Lemma 3.12. **[60]** *There exists a* $\delta > 0$, *such that for the roots* k_1, k_2 *of (3.22) the following estimates hold*

1. *If* $\kappa < 0$, *then*

$$\begin{aligned} |k_1| &\leq 1-\delta, && for\ |z| \geq 1, \\ |k_2| &> 1, && for\ |z| \geq 1,\ z \neq 1, \\ k_2 &= 1, && for\ z = 1. \end{aligned}$$

2. *If* $\kappa > 0$, *then*

$$\begin{aligned} |k_1| &< 1, && for\ |z| \geq 1,\ z \neq 1, \\ k_1 &= 1, && for\ z = 1, \\ |k_2| &\geq 1+\delta, && for\ |z| \geq 1. \end{aligned}$$

Proof. [of Theorem 3.11] To solve the difference equation (3.20)-(3.21b) for $|z| > 1$, we note that the general solution of (3.20) belonging to $l^2(x)$ has the form

$$\tilde{v}_j = k_1^j \varphi_1,$$

where k_1 is the (smaller) root of the characteristic equation (3.22). We insert this solution into the condition (3.21b) and obtain

$$\varphi_1 (k_1 - 1)^2 = 0.$$

But, according to the previous lemma, $|k_1 - 1| \geq \delta$. Hence, equations (3.20)-(3.21b) have no eigenvalues.
To determine whether $z = 1$ is a generalized eigenvalue of (3.20)-(3.21b), we substitute $z = 1$ into equation (3.22) and obtain

$$k = \frac{\kappa^2 \pm |\kappa|}{\kappa^2 - \kappa} = \begin{cases} 1 & =: k_2 \\ \dfrac{\kappa^2 + \kappa}{\kappa^2 - \kappa} & =: k_1. \end{cases}$$

For this scheme, k_1 is not relevant, since $|k_1| = \left|\dfrac{\kappa^2+\kappa}{\kappa^2-\kappa}\right| < 1$ (for $-1 \leq \kappa < 0$), and hence k_1 will not satisfy equation (3.21b).
We notice that for $|z| > 1$, k_2 will satisfy $|k_2| > 1$. This is the case because k_1 is clearly inside the circle $|z| = 1$, so k_2 must be outside that circle. Since $|z| = 1$ is associated with k_2, the solution at $|z| = 1$, does not satisfy condition 3 of Definition 3.9. Thus, $z = 1$ is not a generalized eigenvalue, and the difference scheme (3.18)-(3.19b) is stable.
We emphasize that we have already assumed that the difference scheme is stable as an initial value problem scheme. Hence, the stability proved here will be conditional stability with condition $-1 \leq \kappa < 0$.
Considering the case (3.19a), we substitute

$$\tilde{v}_j = k_1^j \varphi_1, \qquad |k_1| \leq 1 - \delta,$$

into boundary condition (3.19a)

$$\varphi_1(z - 1 + \kappa k_1 - \kappa) = 0.$$

For $-1 \leq \kappa < 0$ (the stability condition for the Cauchy problem) and $|z| \geq 1$, we have

$$\begin{aligned} |z - 1 + \kappa k_1 - \kappa| &> \left|1 + k_1 \frac{\kappa}{z - 1 - \kappa}\right| \\ &\geq 1 - k_1 \left|\frac{\kappa}{1 + \kappa - z}\right| \\ &\geq 1 + (\delta - 1)\left|\frac{\kappa}{1 + \kappa - z}\right| \geq \delta. \end{aligned}$$

It follows that (3.20)-(3.21a) has no eigenvalues.
Analogue to the computations used in the first part, we show that $z = 1$ is not a generalized eigenvalue of (3.20)-(3.21a) . □

3.3 Numerical tests

In the following numerical experiments we compare the performance of the ABCs and the FBCs, as well as the numerical approximation with FBCs for different scaling matrices.

3.3.1 Example 1

Consider the linear hyperbolic system

$$\begin{pmatrix} u \\ v \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_x + \begin{pmatrix} 1 & 1 \\ \frac{3}{4} & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}, x \in I\!R, \quad (3.23a)$$

$$u(x,0) = u^0(x), \qquad v(x,0) = v^0(x), \quad (3.23b)$$

where u^0, v^0, f and g have compact support in $(0,1)$.
The corresponding steady equation on $I\!R$ is given by

$$\begin{pmatrix} u \\ v \end{pmatrix}_x + \begin{pmatrix} 1 & 1 \\ -\frac{3}{4} & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f \\ -g \end{pmatrix}, \quad x \in I\!R, \quad (3.24a)$$

with the decay condition

$$u, v \to 0, \qquad x \to \pm\infty. \quad (3.24b)$$

The zero and first order ABCs for the restriction to the interval $0 \le x \le 1$ are, respectively

$$u = 0, \qquad x = 0, \quad (3.25a)$$

$$v = 0, \qquad x = 1, \quad (3.25b)$$

and

$$u_t + v/2 = 0, \qquad x = 0, \quad (3.26a)$$

$$v_t + 3u/8 = 0, \qquad x = 1. \quad (3.26b)$$

The matrix

$$S = \begin{pmatrix} a & 2a/3 \\ b & 2b \end{pmatrix}, \qquad 0 \neq a, b \in I\!R,$$

transforms the steady state problem (3.24a) to the diagonal form. Diagonalize $\Lambda^{-1}C$

$$S\Lambda^{-1}CS^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}. \quad (3.27)$$

For the decay condition to be valid we need

$$u + 2v/3 = 0, \qquad x = 0, \quad (3.28a)$$

$$u + 2v = 0, \qquad x = 1. \quad (3.28b)$$

The first order FBCs combine ABCs (3.26) and the steady boundary conditions (3.28), in analogue to (2.17),

$$u_t + a(u + 2v/3) = 0, \qquad x = 0, \tag{3.29a}$$

$$v_t + b(u + 2v) = 0, \qquad x = 1. \tag{3.29b}$$

The finite difference scheme introduced in the first section is used in the following numerical tests.

(i) Consider (3.23) with zero initial condition, $f = 0$, and

$$g(x) = \begin{cases} \cos^2(\pi(x - 0.5)/0.9), & x \in (0.05, 0.95), \\ 0, & elsewhere, \end{cases}$$

together with each of the boundary conditions (3.26) and (3.29). The convergence as $t \to \infty$ of the solution of the resulting IBVP to the solution of the steady unbounded problem has been tested ($h = 0.0005,\ r = k/h = 0.9$). The steady state solution (3.24) is given in Figure 3.1 and the convergence to this solution is described in Figures 3.2 and 3.3.

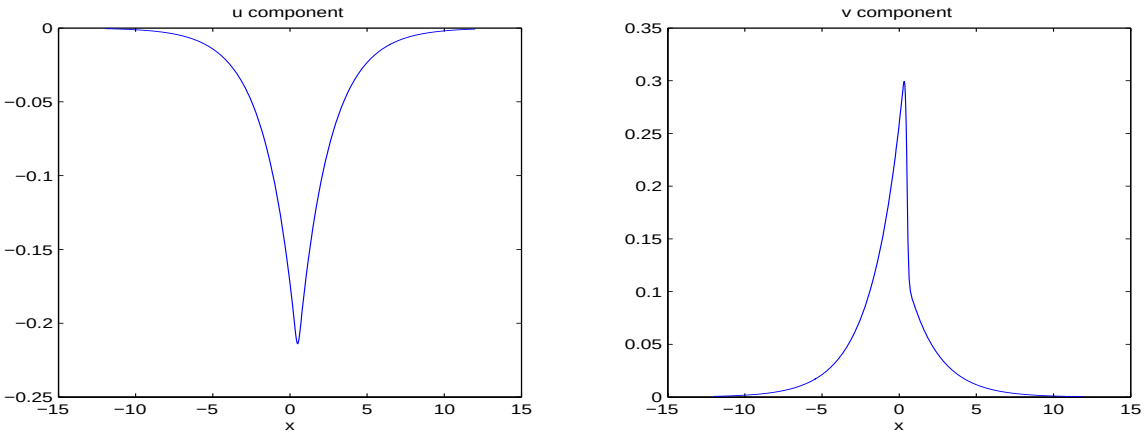


Figure 3.1: Steady state solution of (3.24).

Figure 3.2 shows, with different choices of a and b, that the solution of (3.23) with the new boundary conditions (3.29) converges in $(0, 1)$ to the solution of the steady unbounded problem. Figure 3.3 shows that this is not true for the first order boundary conditions (3.26).

(ii) Using equation (2.20), the optimal choices of a and b are

$$a = \frac{\lambda_2 c_{12}}{\lambda_2 - \lambda_1} \frac{1}{s_{12}} = \frac{3}{4}, \qquad b = \frac{\lambda_1 c_{21}}{\lambda_1 - \lambda_2} \frac{1}{s_{21}} = \frac{3}{8},$$

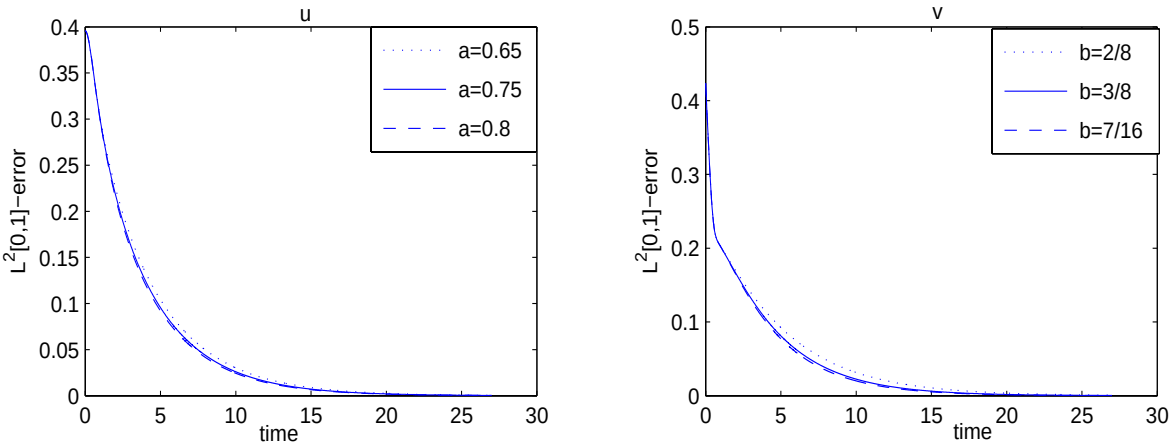


Figure 3.2: $L^2(0,1)-$error between the solution with boundary condition (3.29) and the steady state solution.

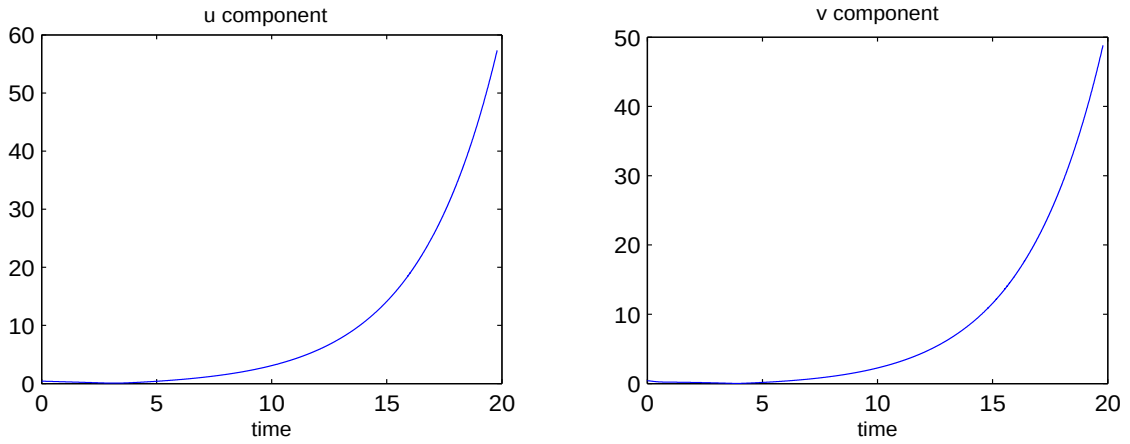


Figure 3.3: $L^2(0,1)-$error between the solution with boundary condition (3.26) and the steady state solution.

and the FBCs (3.29a) become

$$u_t + \frac{3}{4}(u + 2v/3) = 0, \qquad x = 0, \tag{3.30a}$$

$$v_t + \frac{3}{8}(u + 2v) = 0, \qquad x = 1. \tag{3.30b}$$

In this test we show that this choice, among other arbitrary choices, improve the approximation for short time computations. The convergence to steady state is tested, for arbitrary non-zero constants a and b, in part (i). Therefore, for short time comparison, it is reasonable to consider $f(x) = g(x) = 0$.
Since an asymptotic approximation is used to localize the exact nonlocal

boundary conditions, we consider a highly-oscillatory initial data (2.7).

$$u(x,0) = v(x,0) = \begin{cases} \cos^2(2\pi(x-0.5))\sin(2\pi p x), & x \in (0.25, 0.75), \\ 0, & elsewhere, \end{cases} \tag{3.31}$$

The cases $p = 10, 20$ are described in Figure 3.4.

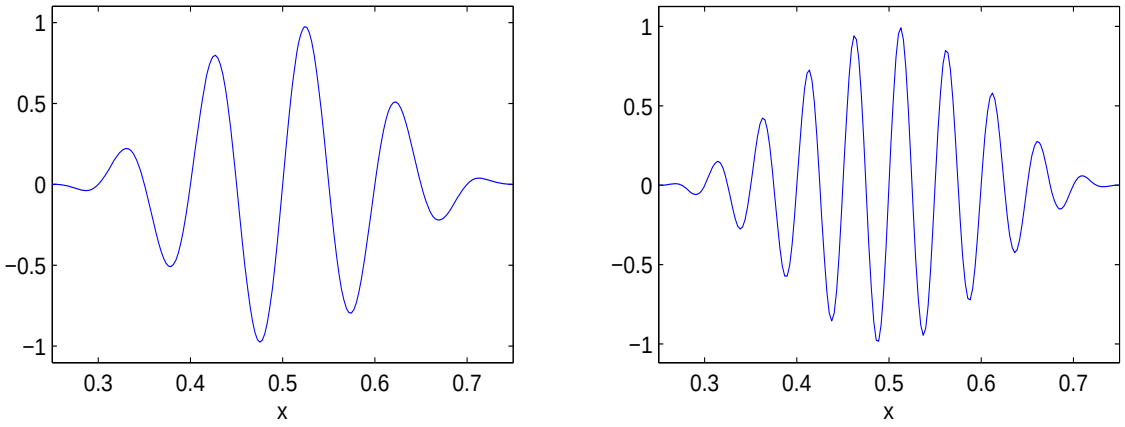


Figure 3.4: Initial values (3.31). Left: $p = 10$. Right: $p = 20$.

The error between the exact solution and the solution with the boundary condition (3.29) for different values of a and b has been tested. The absolute errors of the inflow data (u at $x = 0$ and v at $x = 1$) and the $L^2(0,1)$-error are considered. The step size is chosen small ($h = 0.0005$) in order to estimate

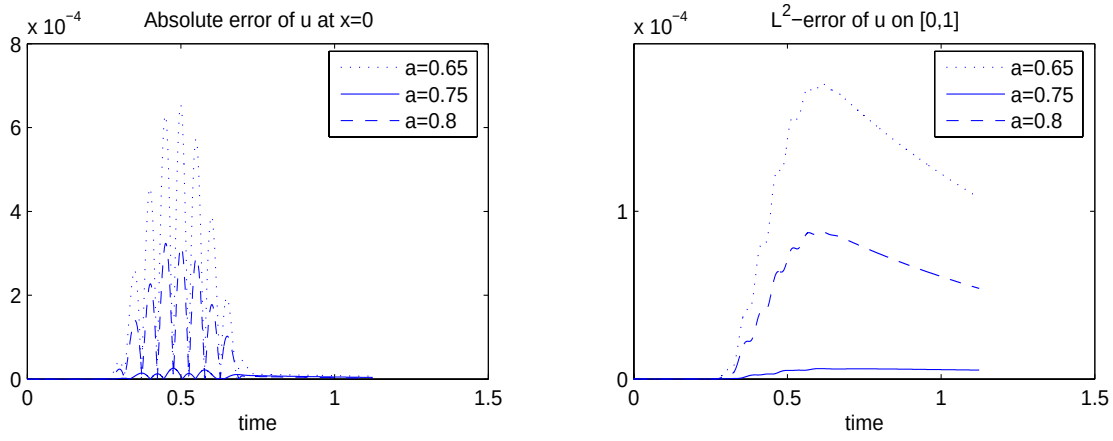


Figure 3.5: Comparison of the error between the exact solution u and the solution with boundary conditions (3.29) for different values of a.

the errors due to the boundary conditions and not the discretization errors. In

case $p = 10$, Figures 3.5-3.6 clearly show that our choices of a and b give the minimum error. The same result holds for the case $p = 20$.

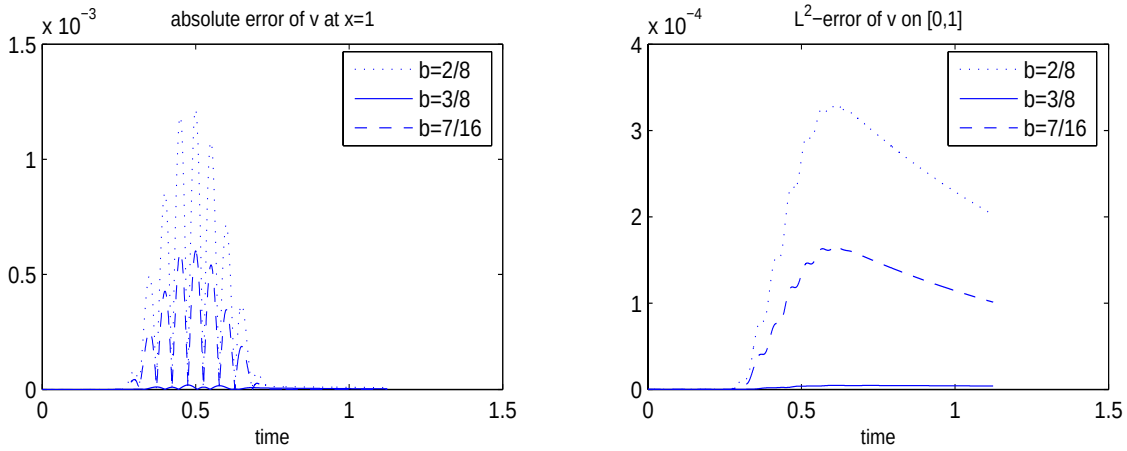


Figure 3.6: Comparison of the error between the exact solution v and the solution with boundary conditions (3.29) for different value of b.

(iii) In this example, we test the dependence of the boundary condition on the initial frequency. We consider the system (3.23), boundary conditions (3.30), and the initial data (3.31). In the case $p = 10$, Figure 3.7 compares the absolute error of inflow data for different refinements of the space step size h. Figures 3.8-3.9 show the same comparison but for the case $p = 20$.

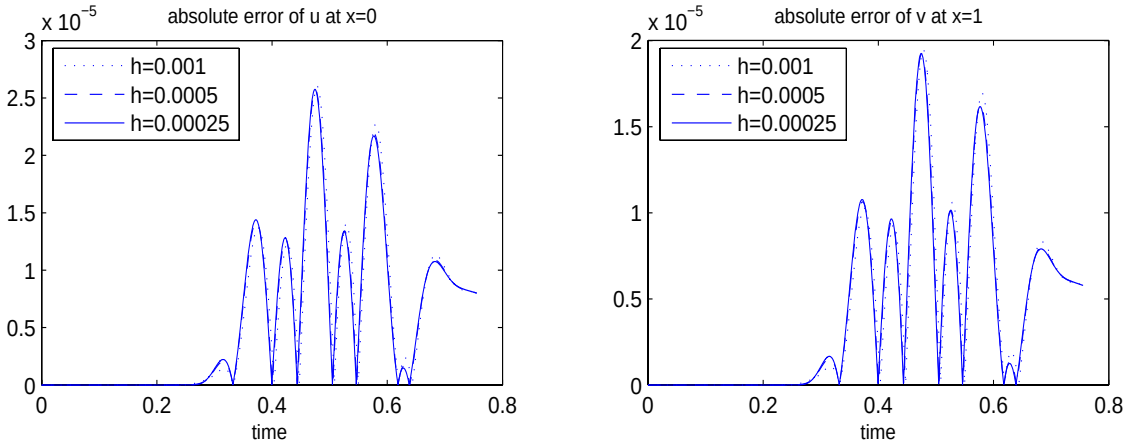


Figure 3.7: Comparison of errors between the exact solution and the solution with the boundary conditions (3.30) for different h, $p = 10$.

Tables 3.1-3.2 list the maximal absolute errors at the inflow data, the tables show that as h is getting smaller the error is reduced slower. As a result the

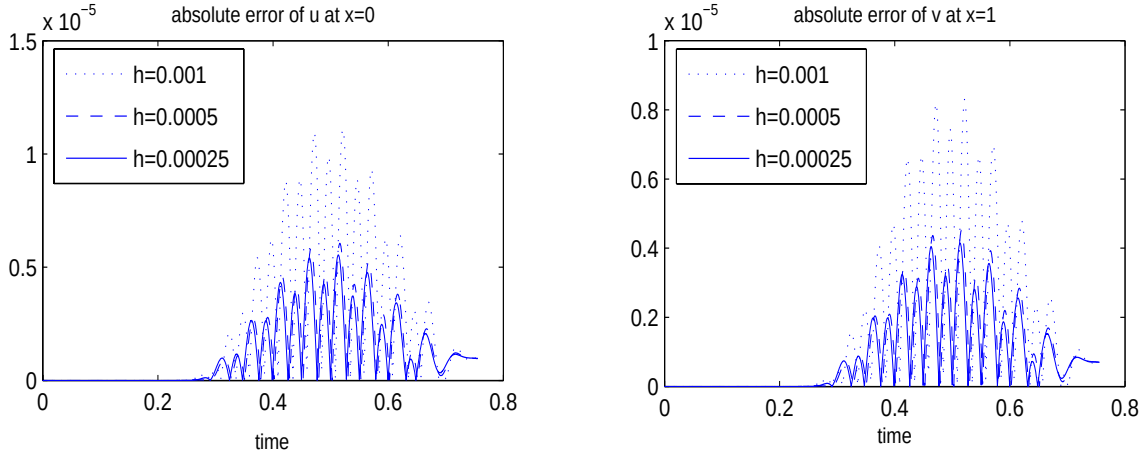


Figure 3.8: Comparison of errors between the exact solution and the solution with the boundary conditions (3.30) for different h, $p = 20$.

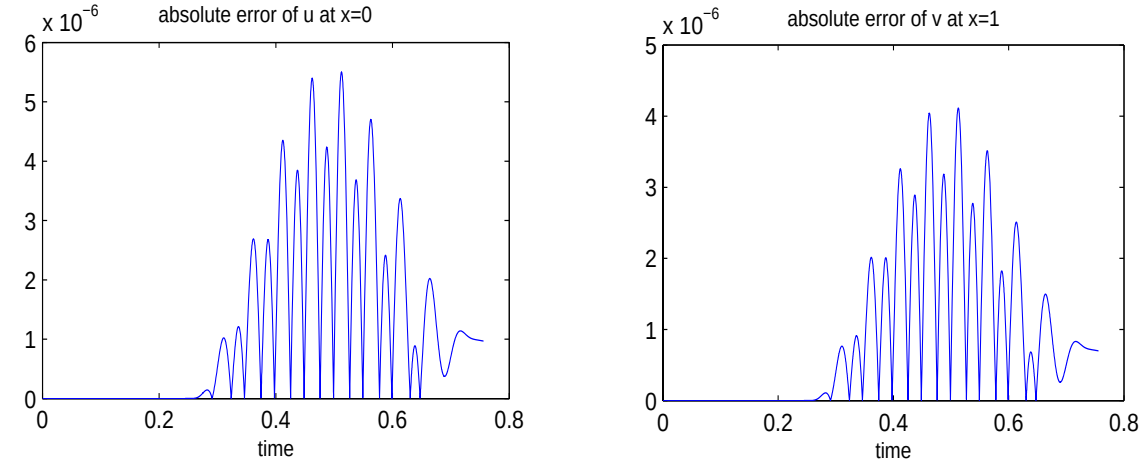


Figure 3.9: Comparison of errors between the exact solution and the solution with the boundary conditions (3.30) for $h = 0.0001$ and $p = 20$.

last row of the two tables are good approximations of the errors due to the boundary conditions (3.30). The maximal absolute error for the case $p = 10$ of u at $x = 0$ and v at $x = 1$ are $2.5702 \cdot 10^{-5}$ and $1.9215 \cdot 10^{-5}$, respectively. In the case of $p = 20$, they are reduced to $5.5009 \cdot 10^{-6}$ and $4.1164 \cdot 10^{-6}$, respectively. This shows that with highly oscillating initial data the errors become smaller, which agrees with the approximation of the nonlocal exact boundary condition (2.7) with asymptotic expansion.

h	u at $x=0$, $(\cdot 10^{-5})$	v at $x=1$, $(\cdot 10^{-5})$
0.001	2.5997	1.9439
0.0005	2.5713	1.9227
0.00025	2.5702	1.9215

Table 3.1: Maximal absolute error, $p = 10$.

h	u at $x=0$, $(\cdot 10^{-6})$	v at $x=1$, $(\cdot 10^{-6})$
0.001	11.079	8.2997
0.0005	6.0525	4.5299
0.00025	5.5394	4.1452
0.0001	5.5009	4.1164

Table 3.2: Maximal absolute error, $p = 20$.

Convergence to the steady state

Using the Laplace transform approach, we want to prove the convergence of the IBVP in Example one to the steady state solution. Namely, we consider

$$\begin{pmatrix} u \\ v \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_x + \begin{pmatrix} 1 & 1 \\ \frac{3}{4} & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}, 0 < x < 1, \quad (3.32a)$$

$$u(x,0) = u^0(x), \qquad v(x,0) = v^0(x), \quad (3.32b)$$

with the first order FBCs

$$u_t + \frac{3}{4}(u + 2v/3) = 0, \quad x = 0, \quad (3.32c)$$

$$v_t + \frac{3}{8}(u + 2v) = 0, \quad x = 1, \quad (3.32d)$$

The corresponding steady state equation on $(0,1)$

$$\begin{pmatrix} u^* \\ v^* \end{pmatrix}_x + \begin{pmatrix} 1 & 1 \\ -\frac{3}{4} & -1 \end{pmatrix} \begin{pmatrix} u^* \\ v^* \end{pmatrix} = \begin{pmatrix} f(x) \\ -g(x) \end{pmatrix}, \quad 0 < x < 1, \quad (3.33a)$$

and the transparent boundary conditions

$$u^* + 2v^*/3 = 0, \qquad x = 0, \quad (3.33b)$$

$$u^* + 2v^* = 0, \qquad x = 1. \quad (3.33c)$$

Define v as

$$v(x,t) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} := \begin{pmatrix} u - u^* \\ v - v^* \end{pmatrix}.$$

Then, v satisfies

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}_x + \begin{pmatrix} 1 & 1 \\ \frac{3}{4} & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0, \quad 0 < x < 1, \tag{3.34a}$$

$$v(x,0) = v^0(x) = \begin{pmatrix} v_1^0(x) \\ v_2^0(x) \end{pmatrix} = \begin{pmatrix} u^0(x) - u^*(x) \\ v^0(x) - v^*(x) \end{pmatrix} \tag{3.34b}$$

$$v_{1,t} + \frac{3}{4}v_1 + \frac{1}{2}v_2 = 0, \qquad x = 0, \tag{3.34c}$$

$$v_{2,t} + \frac{3}{8}v_1 + \frac{3}{4}v_2 = 0, \qquad x = 1. \tag{3.34d}$$

We write

$$\tilde{v}(x,s) = \int_0^\infty e^{-st} v(x,t)dt, \quad s = \alpha + i\tau \quad \tau, \alpha \in I\!R, \quad \alpha > 0,$$

to denote Laplace transform in t. Taking the Laplace transform of (3.34), we obtain

$$\begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix}_x = \begin{pmatrix} -(1+s) & -1 \\ \frac{3}{4} & 1+s \end{pmatrix} \begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix} + \begin{pmatrix} v_1^0(x) \\ -v_2^0(x) \end{pmatrix}, \quad 0 < x < 1, \tag{3.35a}$$

$$s\tilde{v}_1 + \frac{3}{4}\tilde{v}_1 + \frac{1}{2}\tilde{v}_2 = v_1^0(0), \qquad x = 0, \tag{3.35b}$$

$$s\tilde{v}_2 + \frac{3}{8}\tilde{v}_1 + \frac{3}{4}\tilde{v}_2 = v_2^0(1), \qquad x = 1. \tag{3.35c}$$

Since $\Re s > 0$, (3.35a) has the two eigenvalues, η and $-\eta$, where

$$\eta = \sqrt{(1+s)^2 - 3/4},$$

and $\Re\eta > 0$.
The eigenvectors of $\eta, -\eta$ read respectively

$$\begin{pmatrix} -1 \\ \eta + s + 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 \\ \eta - s - 1 \end{pmatrix}.$$

The solution of (3.35) is written as a sum of homogeneous and particular solutions

$$\begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix} = \begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix}_h + \begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix}_p, \tag{3.36}$$

where

$$\begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix}_h = \sigma_1 \begin{pmatrix} -1 \\ \eta+s+1 \end{pmatrix} e^{\eta x} + \sigma_2 \begin{pmatrix} 1 \\ \eta-s-1 \end{pmatrix} e^{-\eta x}, \tag{3.37}$$

and $v_p(x,s) = \begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix}_p$ is any particular solution.

Using variation of parameters, v_p can be found as

$$v_p(x,s) = \psi(x,s) \int_0^x \psi^{-1}(y,s) v^0(y) dy,$$

where

$$\psi(x,s) = \begin{pmatrix} -e^{\eta x} & e^{-\eta x} \\ (\eta+s+1)e^{\eta x} & (\eta-s-1)e^{-\eta x} \end{pmatrix},$$

is the fundamental solution. This can be simplified and written finally as

$$\begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix}_p = \int_0^x \begin{pmatrix} v_1^0(y)cosh(\eta(x-y)) - \frac{1}{\eta}((s+1)v_1^0(y) + v_2^0(y))sinh(\eta(x-y)) \\ v_2^0(y)cosh(\eta(x-y)) + \frac{1}{\eta}(\frac{3}{4}v_1^0(y) + (s+1)v_2^0(y))sinh(\eta(x-y)) \end{pmatrix} dy \tag{3.38}$$

Applying boundary conditions (3.35b)-(3.35c) to (3.36) gives

$$B(s) \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} = \begin{pmatrix} v_1^0(0) \\ h(s) \end{pmatrix}, \tag{3.39}$$

where $B(s)$ is given by

$$B(s) = \begin{pmatrix} \eta-s-1/2 & \eta+s+1/2 \\ [(s+3/4)(\eta+s+1)-3/8]\, e^{\eta} & [(s+3/4)(\eta-s-1)+3/8]\, e^{-\eta} \end{pmatrix}, \tag{3.40}$$

and

$$h(s) := -(s+3/4)\tilde{v}_{2p}(1) - 3\tilde{v}_{1p}(1)/4 + v_1^0(1).$$

The equation (3.39) has a unique solution if and only if the determinant of $B(s)$ is nonzero.

Denote the determinant of $B(s)$ by $d_B(s)$, then

$$d_B(s) = (4s^3 + 10s^2 + 6s + 3/4)\sinh(\eta) + (4s^2 + 6s + 3/2)\cosh(\eta).$$

Now,

$$d_B(s) = 0, \tag{3.41}$$

has no solution with positive real part (Figure 3.10), and it has only real negative roots (Figure 3.11). Using (3.39) we solve for σ_1, σ_2, and substitute the result in

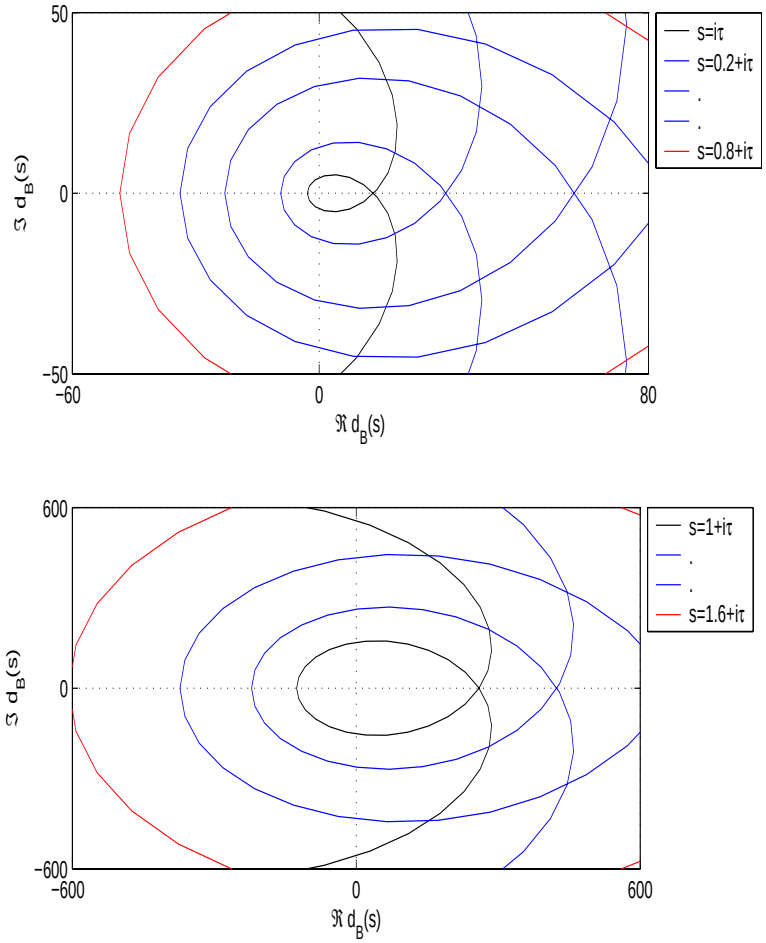


Figure 3.10: Contour plots of $d_B(s)$ for $\Re s \times \Im s \in [0, 1.6] \times [-10, 10]$.

(3.37) to obtain

$$\begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix} = \frac{1}{d_B(s)} \left[\begin{pmatrix} -1 \\ \eta + s + 1 \end{pmatrix} h_1(s) e^{\eta} + \begin{pmatrix} 1 \\ \eta - s - 1 \end{pmatrix} h_2(s) e^{-\eta} \right] + \begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix}_p \tag{3.42}$$

where

$$h_1(s) := [(s + 3/4)(\eta - s - 1) + 3/8] v_1^0(0) e^{-\eta} - (\eta + s + 1/2) h(s)],$$

$$h_2(s) := [-(s + 3/4)(\eta + s + 1) + 3/8] v_1^0(0) e^{\eta} - (\eta - s - 1/2) h(s)],$$

and the particular solution is given by (3.38).
Using the inversion formula,

$$v(x,t) = \frac{1}{2\pi i}\int_{\beta-i\infty}^{\beta+i\infty} \tilde{v}(x,s)e^{st}ds,$$

where β is larger than the real part of any pole of the integrand. The poles of the first part of (3.42) are simply the zeros of the $d_B(s)$, which are simple poles and have negative real part, while the second part has no poles. According to the complex inversion theorem [**21**]

$$\begin{aligned}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= \sum \textit{residues of } e^{st}\begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{pmatrix}_h \textit{ at each of its singularities in } \mathbb{C} \\ &= \sum_{s=s_i}\left[\begin{pmatrix} -1 \\ \eta+s+1 \end{pmatrix} h_1(s)e^{\eta} + \begin{pmatrix} 1 \\ \eta-s-1 \end{pmatrix} h_2(s)e^{-\eta}\right]\frac{e^{st}}{d_B'(s)}\end{aligned}$$

where $s_i \in I\!R^-$ (Figure 3.11). Hence, v converges to 0 as $t \to \infty$.

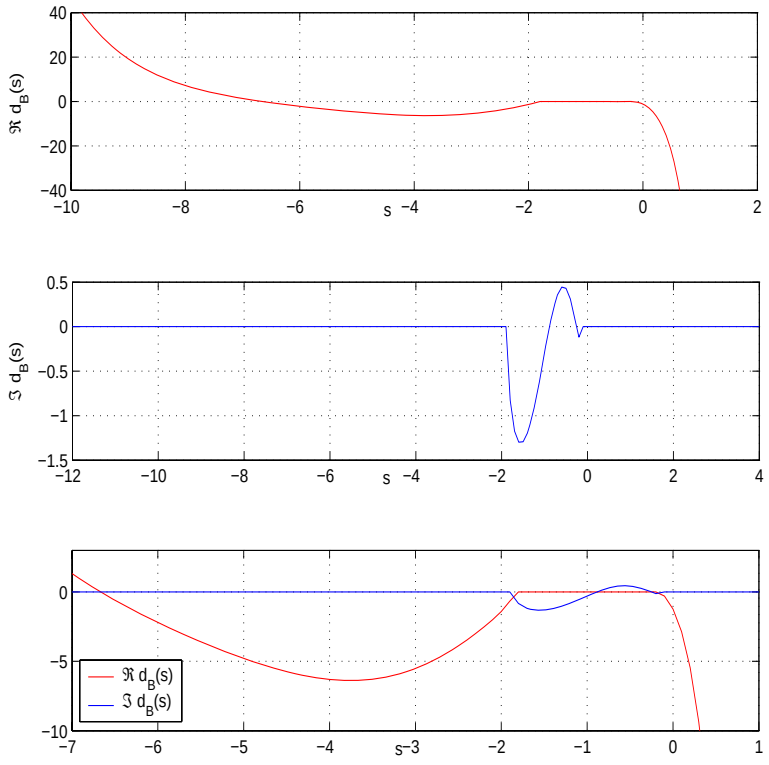


Figure 3.11: Graph of $d_B(s)$, $s \in I\!R$.

3.3.2 Example 2

Consider (3.1a) in $\mathbb{R}$ with

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & -0.8 \end{pmatrix}, \quad C = \begin{pmatrix} 0.2 & 0 & 1 \\ 0 & 0.4 & 2 \\ -1 & -2 & 1 \end{pmatrix}, \tag{3.43}$$

and $f(x,t) = (f_1(x), f_1(x), f_1(x))^T$, where

$$f_1(x) = \begin{cases} 10\exp(-100(2x-1)^2), & x \in (0.25, 0.75), \\ 0, & elsewhere. \end{cases} \tag{3.44}$$

The initial function is given by (Figure 3.12)

$$u^0(x) = \begin{cases} \cos(\pi(x-0.5)/0.9), & x \in (0.05, 0.95), \\ 0, & elsewhere. \end{cases} \tag{3.45}$$

In the first part of this example we compare the performance of the FBCs and

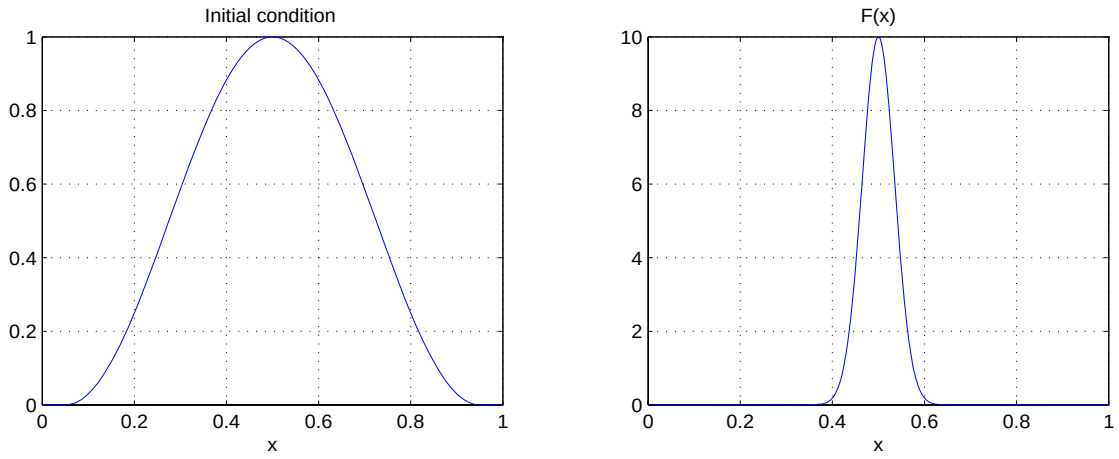


Figure 3.12: Left: Initial condition (3.45). Right: Forcing function (3.44).

ABCs for long time computations. The first order ABCs for the restriction to the interval $0 \le x \le 1$ are

$$u_{1,t} + 0.444u_{3,t} = 0, \quad x = 0, \tag{3.46a}$$

$$u_{2,t} + 1.6u_{3,t} = 0, \quad x = 0, \tag{3.46b}$$

$$u_{3,t} - 0.5556u_{1,t} - 0.4u_{2,t} = 0, \quad x = 1, \tag{3.46c}$$

while the first order FBCs read

$$u_t^+ + V^+(S^{++}u^+ + S^{+-}u^-) = 0, \qquad x = 0, \tag{3.47a}$$
$$u_t^- + V^-(S^{-+}u^+ + S^{--}u^-) = 0, \qquad x = 1. \tag{3.47b}$$

The matrix S, which diagonalizes $\Lambda^{-1}C$, is given by

$$S = \begin{pmatrix} -0.2287 & -0.6791 & -1.0083 \\ 1.0522 & -0.0952 & 0.0656 \\ -0.2690 & -0.4004 & 1.1275 \end{pmatrix}.$$

The scaling matrices

$$V^+ = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad V^- = e$$

are chosen as a general solution of $V^+S^{+-} = X_1^{+-}$ and $V^-S^{-+} = X_1^{-+}$, respectively. Following the procedure presented in Section 1.2, we take

$$V^{*+} = \begin{pmatrix} -0.4389 & 0.0285 \\ -1.5801 & 0.1028 \end{pmatrix}, \qquad V^{*-} = 1.3306.$$

The stepsizes are chosen small in order to see the errors due to different boundary conditions and not the discretization errors ($h = 0.0005$, $k = 0.0004$). It is clear that the CFL-condition, $\max_{j=1,2,3} |r\lambda_j| < 1$, is satisfied.
The steady state solution is given in Figure 3.13-Right and the convergence to this solution as $t \to \infty$ in $(0, 1)$ is described in Figure 3.14. In Figures 3.15-3.17 the solutions with FBCs for different choices of the scaling matrices are compared to the exact solution over $(-\infty, \infty)$. The plots show that the FBCs with the proposed optimal choices of V^+, V^- give the best approximate solutions in the transient phase to the exact solution in the unbounded domain.
Tables 3.3-3.5 list the maximal absolute errors at the inflow data (u_1, u_2 at $x = 0$ and u_3 at $x = 1$). As well as the $L^2(0, 1)-$error between exact solution and the solution with the boundary condition (3.29) for different values of a, b, c, d, and e.
The numerical results give quantitative evidence that the FBCs are useful for both short and long times.

a,	b	abs. error at $x = 0$	$L^2(0,1)-$error
-0.2,	0	0.0804	0.0548
-0.8,	0.2	0.0829	0.0506
a^*,	b^*	0.0318	0.0293

Table 3.3: Maximum errors due to the first order FBCs of u_1 for different choices of a and b, see also Figure 3.15.

c,	d	abs. error at $x = 0$	$L^2(0,1)-$error
-1,	0	0.2208	0.0664
-2.5,	0.5	0.1312	0.0430
c^*,	d^*	0.1233	0.0408

Table 3.4: Maximum errors of u_2 due to the first order FBCs for different choices of c and d, see also Figure 3.16.

e	abs. error at $x = 1$	$L^2(0,1)-$error
1	0.1080	0.0421
2.5	0.0943	0.0468
e^*	0.0875	0.0344

Table 3.5: Maximum errors due to the first order FBCs of u_3 for different choices of e, see also Figure 3.17.

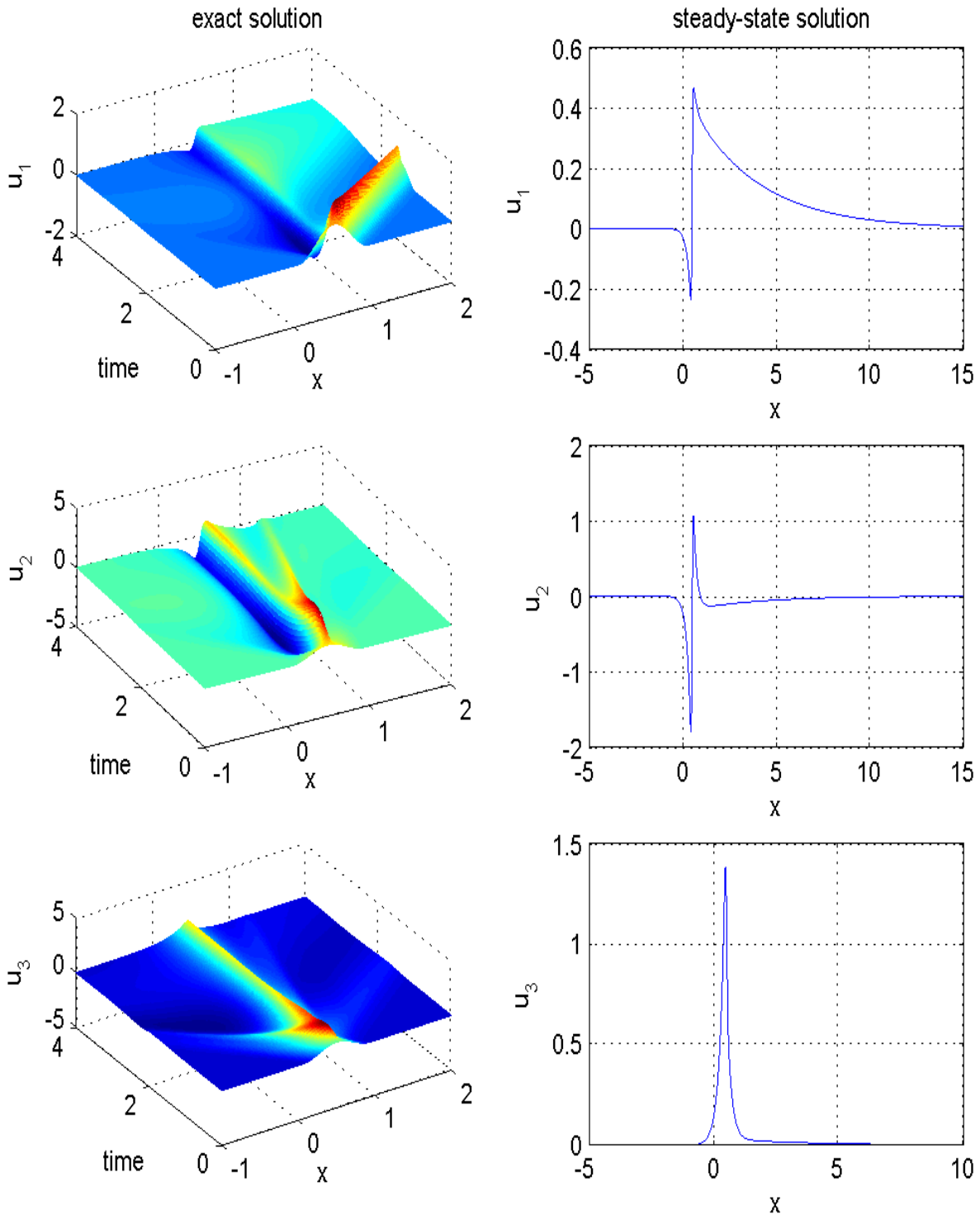


Figure 3.13: Left: Exact solution. Right: Steady state solution.

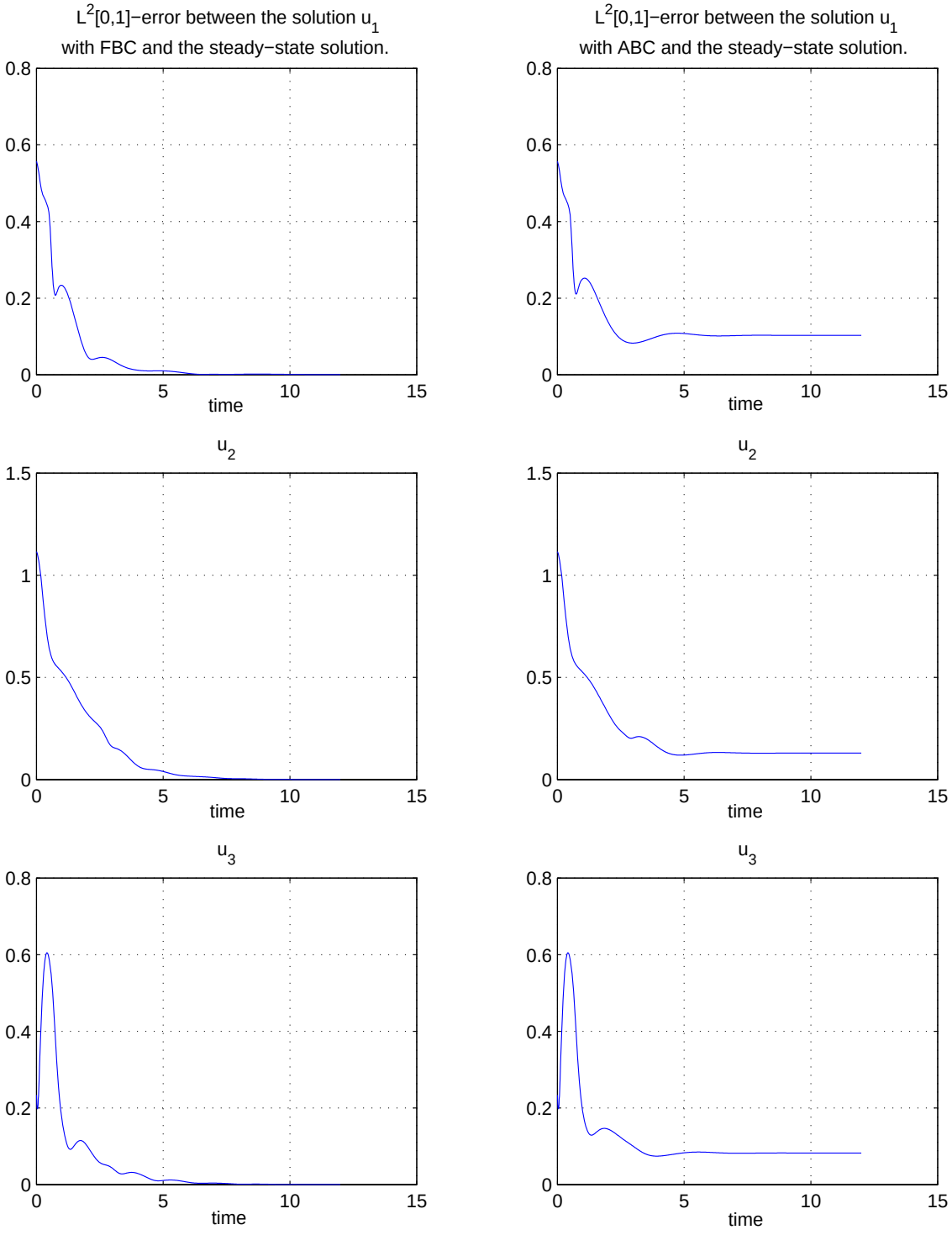


Figure 3.14: Convergence to the steady state solution as $t \to \infty$ in $(0, 1)$.

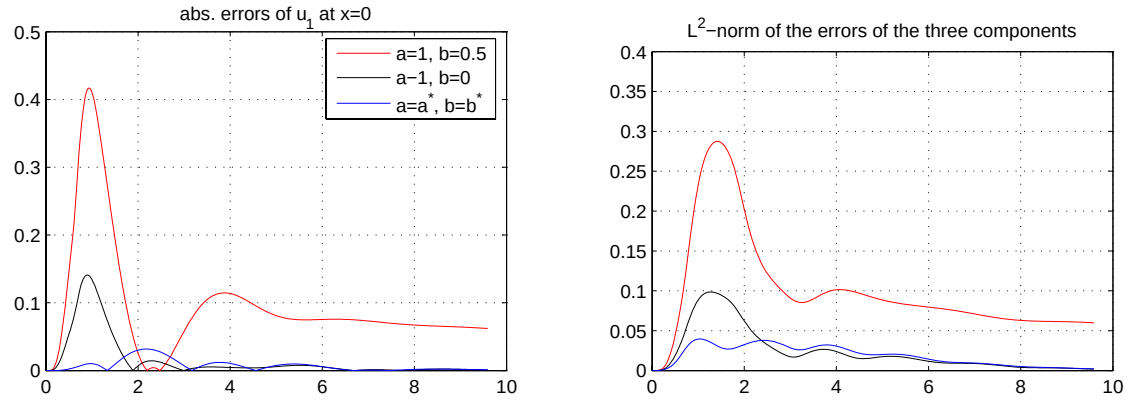


Figure 3.15: Comparison of the errors between the exact solution of u_1 and the solution with FBCs for different choices of a and b.

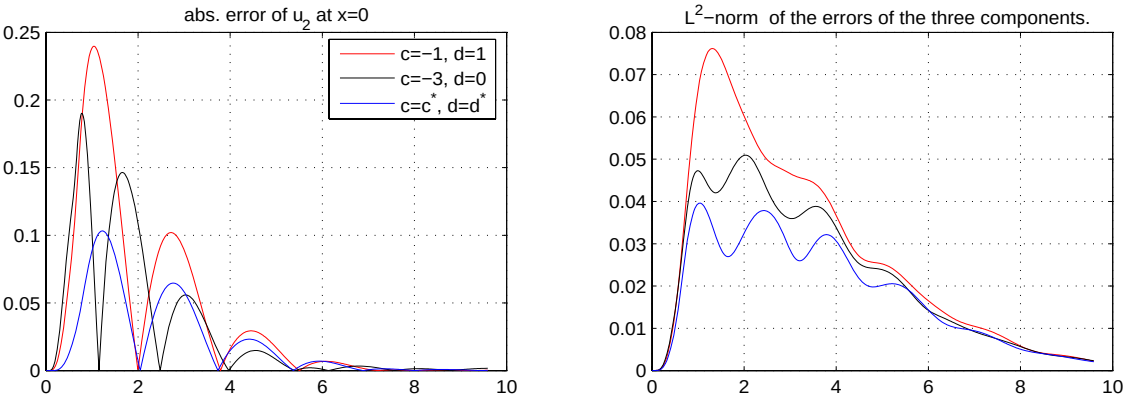


Figure 3.16: Comparison of the errors between the exact solution of u_2 and the solution with FBCs for different choices of c and d.

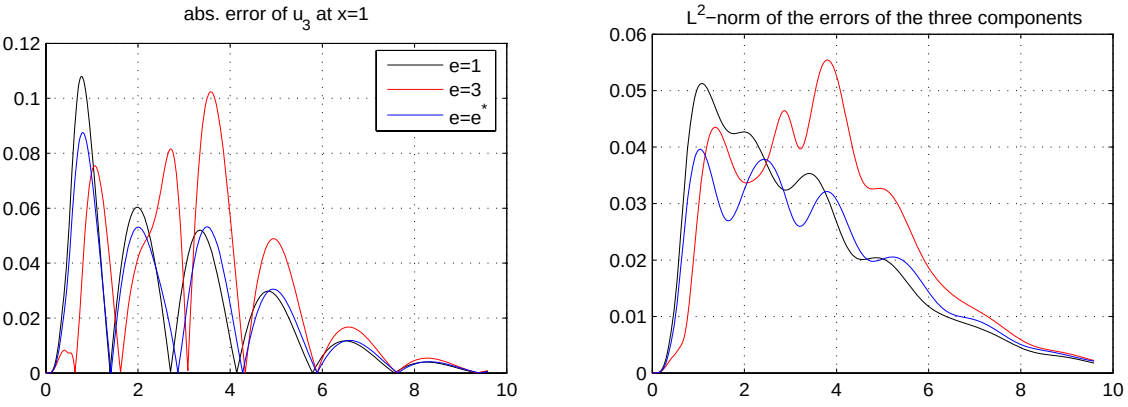


Figure 3.17: Comparison of the errors between the exact solution and the solution with FBCs for different choices of e.

Part II

PML Absorbing Boundary Condition

Chapter 4

Numerical solution to nonlinear KG equations by PML approach

Abstract

It is a common practice to use absorbing boundary conditions when simulating waves in unbounded domains. In this paper, we propose a perfectly matched layer (PML) approach for the numerical solution to nonlinear Klein-Gordon (KG) equations. The procedure includes four steps: Firstly, the nonlinear KG equation is transformed into a semi-linear hyperbolic system with a damping term by introducing auxiliary unknown functions. Secondly, we linearize the damping term and design a PML formulation for the linearized system. Then, we derive a nonlinear PML system by replacing the linearized damping term with its original nonlinear counterpart. Finally, an implicit-explicit finite difference scheme is used to solve the nonlinear PML system. This approach is next extended to the two-dimensional case. The numerical tests show the efficiency of this "PML linearization" over other local absorbing boundary conditions.

4.1 Introduction

[1] Waves in unbounded domains exist in a wide range of application areas, such as quantum mechanics, aeroacoustics, electromagnetics, fluid dynamics, and geophysics [**29**]. When seeking numerical solutions of the governing partial differential

[1] The content of this chapter is a joint work with my Ph.D. adviser A. Arnold and C. Zheng [**4**]

equations, it is a common practice to use absorbing boundary conditions (ABCs) to limit the computational domain to a finite region. An appropriate ABC should have the following three features: It should define together with the interior differential equation a well-posed (initial boundary value) problem. Its solution should be a "reasonably accurate" approximation of the original whole space solution. Lastly, an ABC should allow for an efficient numerical implementation.

Most ABCs in the literature can be classified into PDE-based and material-based. PDE-based ABCs are imposed on some specific artificial boundary. For wave-like equations, they are obtained by factorizing the governing equation into incoming and outgoing modes, while "minimizing" reflections of the outgoing waves. The readers are referred to [**78, 40, 31, 53, 5**] for detailed reviews. Material-based ABCs follow a different philosophy. Instead of using artificial boundaries to limit the computational domain, they use a lossy medium surrounding the physically interesting part of the domain to annihilate (or at least damp) the outgoing waves [**42**].

Very often, methods based on the pseudodifferential operator theory are used to design approximate PDE-based ABCs. Engquist and Majda [**24**] were the first to apply this technique to the numerical simulation of waves. Most recent approaches in this direction include [**73, 74**] by Szeftel. There, he derived a hierarchy of local ABCs for semilinear wave equations and nonlinear Schrödinger equations based on the pseudodifferential and paradifferential calculus.

Material-based ABCs have been used to deal with wave problems for three decades. In early versions, dissipative or damping terms were added into the governing equations [**58, 59**] to form an absorbing layer. Later, purely numerical means have been used to achieve the attenuation of waves. Rai and Morin [**69**], Colonius *et al.* [**19**] created a sponge layer by the grid stretching technique. Hu and Atkins [**56**] pointed out that the stretching has to be performed gradually, otherwise significant reflection can be induced. In 1994, Bérenger [**13**] proposed his famous perfectly matched layer (PML) for computational electromagnetics. The advantage of the PML lies in the fact that the absorbing layer is theoretically reflectionless for multi-dimensional linear waves of any angle and any frequency. As a result, the zone of the PML is usually thin compared with that of other absorbing layer techniques.

The original PML technique of Bérenger was based on the split physical variables that was shown to be only weakly well-posed [**1**]. Later studies revealed that the PML can be considered as a coordinate stretching from the real space to the complex plane [**15, 79, 16, 17**] and a well-posed unsplit PML can be achieved. Extensive

work on the applications of the PML technique has been done, however, only recent studies extended it to some nonlinear problems. For example, Hu [**55**] applied this technique for the nonlinear Euler equations, and Navon *et al.* [**66**] for the nonlinear shallow water equations.

In this paper, we shall apply the PML technique to the nonlinear Klein-Gordon (KG) equation of the form

$$u_{tt} = c^2 \Delta u - \varphi(u, u_t, \nabla u), \tag{4.1}$$

where φ is a given scalar function. This equation has many applications in science and engineering. The linear KG equation plays a fundamental role in quantum field theory [**80**], and it models the motion of a massive spinless particle in the relativistic regime. In the study of semi-conductor devices, the sine-Gordon equation (i.e. (4.1) with $\varphi := \sin u$) is commonly used to model the propagation of fluxons in Josephson junctions. Furthermore, it is used to study the motion of rigid pendula attached to a stretched wire and dislocations in crystals [**20, 22**]. Well-posedness of the Cauchy problem for the nonlinear KG equation was studied in [**65**].

For the linear KG equation, Givoli and Patlashenko [**33**] designed a series of high-order local ABCs. Later, Givoli and Neta [**32**] modified these ABCs into a form with only low-order derivatives by introducing auxiliary unknown functions. Zahim and Guddati [**81**] applied their continued fraction ABC to this problem. Using some special padding elements, they reduced the reflection of evanescent waves significantly. For the one-dimensional sine-Gordon equation, Zheng [**82**] revised the generalized Dirichlet-to-Neumann mapping proposed by Fokas [**27**] and designed a suitable numerical scheme. This approach uses inverse scattering theory and it relies on the full integrability of the sine-Gordon equation. Hence, the extension of this technique to higher dimensions and other nonlinear KG equations seems impossible.

The main goal of this paper is to apply the PML approach to the nonlinear KG equation. To this end (4.1) is first linearized and transformed into a hyperbolic system with a linear damping term. The standard PML technique is then used for this linearized system. A nonlinear PML system is then obtained by replacing the damping term in the PML formulation with its original nonlinear counterpart.

This paper is organized as follows: In Section 2 we illustrate the PML for the one-dimensional KG equation, both for linear and nonlinear cases. In Section 3 we extend this technique to higher dimensional cases. An implicit-explicit scheme is used for solving the nonlinear PML system in Section 4. In Section 4.5 we present some

numerical tests to demonstrate the effectiveness of our approach. We conclude in Section 6.

4.2 One-dimensional KG equations

We start with a linear KG equation of the form

$$u_{tt} = c^2 u_{xx} - Lu,\ x > 0,\quad t > 0, \tag{4.2}$$

$$u(0,t) = B(t),\ u(x,0) = u_0(x),\ u_t(x,0) = u_1(x). \tag{4.3}$$

Here, the constant $c > 0$ is the wave speed and $L \geq 0$ is the dispersion parameter. Further we assume that the initial functions $u_0 \in C^1[0,\infty)$ and $u_1 \in C[0,\infty)$ are compactly supported in $[0,a]$, with some fixed $a > 0$. To assure continuity of the solution to (4.2)-(4.3) we assume $B \in C^1[0,\infty)$ and $B(0) = 0$.
First we transform (4.2)-(4.3) into a hyperbolic system with damping term. Let $v := u_x$ and $w := u_t$. Then (4.2)-(4.3) is equivalent to

$$u_t = w,\ v_t = w_x,\ w_t = c^2 v_x - Lu,\ x > 0,\quad t > 0, \tag{4.4}$$

$$w(0,t) = B'(t),\ u(x,0) = u_0(x),\ v(x,0) = u_0'(x),\ w(x,0) = u_1(x). \tag{4.5}$$

By taking the Laplace transform in t of (4.4), we obtain on the domain $[a,\infty)$

$$s\hat{u} = \hat{w},\ s\hat{v} = \hat{w}_x,\ s\hat{w} = c^2\hat{v}_x - L\hat{u},\ x \geq a. \tag{4.6}$$

Here, $s \in \mathbb{C}$ with $\Re s > 0$ is the dual variable of t. For given s, the bounded modal solution of (4.6) is

$$\hat{u}(x,s) = e^{\lambda x + st},\ \hat{v}(x,s) = \lambda\, e^{\lambda x + st},\ \hat{w}(x,s) = s\, e^{\lambda x + st},\quad x \geq a, \tag{4.7}$$

with arbitrary $t \geq 0$ and

$$\lambda := -\frac{\sqrt[+]{s^2 + L}}{c}.$$

Here, $\sqrt[+]{z}$ denotes the branch of the square with $\Re(\sqrt[+]{z}) > 0$.
The idea of the PML method is to replace (4.4) by an equivalent whole space problem, whose solution decays outside the computational domain (i.e. on $[a,\infty)$) faster (i.e. with a higher exponential rate) than (4.7). Eventually, this new problem will be cut-off at a finite distance and "closed" with an appropriate boundary condition (Dirichlet, e.g.). The PML can also be interpreted as an extension of the independent

real coordinate x to the complex plane [**15, 79, 16**]. More precisely, we introduce the complex change of variables:

$$x' = x'(x) := x + \frac{1}{s+\alpha}\int_a^x \sigma(r)dr, \quad x \geq 0, \tag{4.8}$$

where the phase shift parameter $\alpha \geq 0$ and the absorption function $\sigma \geq 0$ vanish in $[0, a]$. Then, we define modified modal functions as $\hat{u}^m(x,s) := \hat{u}(x'(x), s)$, $\hat{v}^m(x,s) := \hat{v}(x'(x), s)$, $\hat{w}^m(x,s) := \hat{w}(x'(x), s)$ for $x \geq 0$. For $x \geq a$ they read

$$\hat{u}^m(x,s) = e^{\lambda x' + st}, \ \hat{v}^m(x,s) = \lambda\, e^{\lambda x' + st}, \ \hat{w}^m(x,s) = s\, e^{\lambda x' + st}. \tag{4.9}$$

Notice that for different modal solutions (parametrized by s) the corresponding variable transformations (4.8) are different. According to the transformation (4.8), the modified functions $\hat{u}^m$, $\hat{v}^m$, $\hat{w}^m$ coincide with $\hat{u}$, $\hat{v}$, $\hat{w}$ on the x-interval $[0, a]$. However, in $[a, \infty)$ each component is modified by the factor

$$C_f = \exp\left(\frac{\lambda}{s+\alpha}\int_a^x \sigma(r)dr\right) = \exp\left(-\frac{\sqrt[+]{s^2+L}}{c(s+\alpha)}\int_a^x \sigma(r)dr\right),$$

with amplitude

$$|C_f| = \exp\left(-\Re\left(\frac{\sqrt[+]{s^2+L}}{c(s+\alpha)}\right)\int_a^x \sigma(r)dr\right).$$

Apart from the absorption function σ, the damping rate strongly depends on the real part of $\mu := \frac{\sqrt[+]{s^2+L}}{s+\alpha}$. To explain why to introduce a positive phase shift α, we depict in Figure 4.1 $\Re\mu$ when $s \in i\mathbb{R}$

We divide $\mathbb{C} - \left\{\pm\sqrt{L}\, i\right\}$ into two parts: The open line segment $(-\sqrt{L}, \sqrt{L}) \times i$ and its complement. For any s in the first part, $\lambda = -\frac{\sqrt[+]{s^2+L}}{c}$ is a negative real number and the modal solution (4.7) represents an evanescent wave. On the other hand, for any s in the second part, λ is purely imaginary and the modal solution (4.7) represents a traveling wave. Now, compare the two cases $\alpha = 0$ and $\alpha > 0$. When $\alpha = 0$, we always have $\Re\mu > 0$ for the traveling waves, thus the PML technique can indeed damp the modal solution. But for the evanescent waves, we have $\Re\mu = 0$. In this case, although the modal solution itself represents a decaying function, the PML technique cannot improve the decay, and the only effect is a phase transformation of the modal solution. If $\alpha > 0$, then $\Re\mu > 0$ for both the traveling and evanescent waves, see Figure 4.1.

Figure 4.1 also shows that α must not be negative. Otherwise, the modified modal

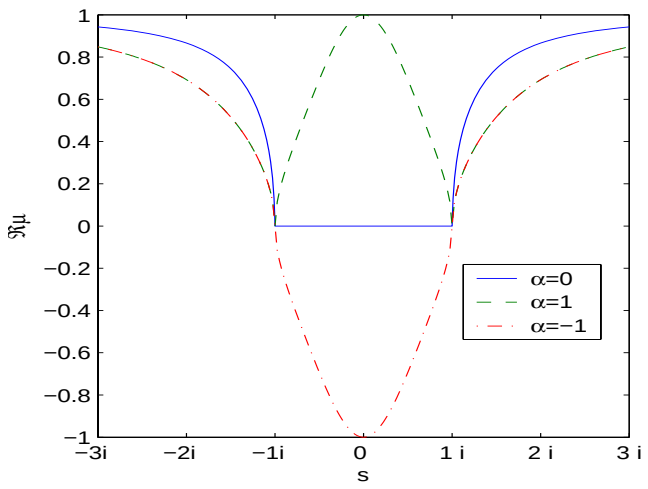


Figure 4.1: Real part of μ for $L = 1$ and $s \in i\mathbb{R}$.

solution could be exponentially increasing as x goes to infinity. A natural question would be how to chose the phase shift α. First we mention that the damping factor $\Re\mu(s;\alpha) \to 1$ for high temporal frequency, i.e. for $\frac{s}{i} \to \pm\infty$. Hence, a natural option would be to minimize $\|1-\Re\mu(s;\alpha)\|_{L^2(\mathbb{R}_{is})}$ w.r.t. $\alpha \geq 0$. But since we are ultimately interested in the *nonlinear* KG equation, we do not elaborate on this issue here.
The exceptional cases $s = \pm\sqrt{L}\,i$ yield $\lambda = 0$. Hence, the modal solution is independent of x and $C_f \equiv 1$. Thus, the PML method would not be able to damp such waves. In the numerical example 1 (cf. §4.5.1) this corresponds to the frequency $\omega = 1$.
Next we seek a PDE system that is satisfied by the modified modal functions (4.9). Using (4.6) and

$$\frac{\partial x'}{\partial x} = \frac{s+\alpha+\sigma(x)}{s+\alpha},$$

we conclude that the modified modal functions satisfy

$$s\hat{u}^m = \hat{w}^m,\ s\hat{v}^m = \frac{s+\alpha}{s+\alpha+\sigma(x)}\hat{w}^m_x,\ s\hat{w}^m = \frac{s+\alpha}{s+\alpha+\sigma(x)}c^2\hat{v}^m_x - L\hat{u}^m \quad \text{on } [a,\infty). \tag{4.10}$$

In order to avoid higher order time derivatives when going back to the time domain, we introduce the auxiliary variables

$$\hat{p}^m(x,s) := \frac{\hat{w}^m_x(x,s)}{s+\alpha+\sigma(x)},\ \hat{q}^m(x,s) := \frac{c^2\hat{v}^m_x(x,s)}{s+\alpha+\sigma(x)}, \qquad x \geq 0.$$

The system (4.10) then becomes

$$s\hat{u}^m = \hat{w}^m, \tag{4.11}$$
$$s\hat{v}^m = \hat{w}_x^m - \sigma(x)\hat{p}^m, \tag{4.12}$$
$$s\hat{w}^m = c^2\hat{v}_x^m - L\hat{u}^m - \sigma(x)\hat{q}^m, \tag{4.13}$$
$$s\hat{p}^m = \hat{w}_x^m - (\alpha + \sigma(x))\hat{p}^m, \tag{4.14}$$
$$s\hat{q}^m = c^2\hat{v}_x^m - (\alpha + \sigma(x))\hat{q}^m. \tag{4.15}$$

Going back to the time domain, we obtain the PML system for the linear KG equation (4.2)

$$u_t^m = w^m, \tag{4.16}$$
$$v_t^m = w_x^m - \sigma p^m, \tag{4.17}$$
$$w_t^m = c^2 v_x^m - Lu^m - \sigma q^m, \tag{4.18}$$
$$p_t^m = w_x^m - (\alpha + \sigma)p^m, \tag{4.19}$$
$$q_t^m = c^2 v_x^m - (\alpha + \sigma)q^m. \tag{4.20}$$

This system was derived for $[a, \infty)$, but we can naturally extend it (setting $\sigma|_{[0,a]} \equiv 0$, $\alpha|_{[0,a]} \equiv 0$) to the whole definition domain $[0, \infty)$ of the original problem (4.2)-(4.3). Since the system (4.16)-(4.20) has only one right traveling characteristic, the only boundary condition is then

$$w^m(0,t) = B'(t), \tag{4.21}$$

and the initial functions are

$$u^m(x,0) = u_0(x),\ v^m(x,0) = u_0'(x),\ w^m(x,0) = u_1(x),\ p^m(x,0) = q^m(x,0) = 0. \tag{4.22}$$

The equations (4.16)-(4.20) constitute a hyperbolic system with damping term. By construction, the solution of (4.16)-(4.20) restricted to $[0, a]$ is exactly the same as that of the original problem (4.2)-(4.3) restricted to $[0, a]$.
Next we consider a more general linear KG equation of the form

$$u_{tt} = c^2 u_{xx} - \varphi(u, u_t, u_x), \quad x > 0,\ t > 0, \tag{4.23}$$

where $\varphi(u, u_t, u_x) := L_1 u + L_2 u_t + L_3 u_x$. The constants are assumed to satisfy $L_1 \geq 0$, $L_2 \geq 0$, and $L_3 \in \mathbb{R}$ which makes φ a damping term. The above PML derivation

can be made analogously, but we omit the details here. The resulting modified PML system reads for $x > 0,\ t > 0$ (dropping the superscript m):

$$u_t = w, \tag{4.24}$$

$$v_t = w_x - \sigma p, \tag{4.25}$$

$$w_t = c^2 v_x - \varphi(u, w, v) - \sigma q, \tag{4.26}$$

$$p_t = w_x - (\alpha + \sigma)p, \tag{4.27}$$

$$q_t = c^2 v_x - (\alpha + \sigma)q. \tag{4.28}$$

As before, $\alpha \geq 0$ and $\sigma \geq 0$ are supported outside the computational domain, i.e. on $[a, \infty)$.

Notice that in our modified PML system (4.24)-(4.28) for the general linear KG equation (4.23), the damping term φ only appears linearly.

Next we consider the KG equation with a nonlinearity φ. In order for φ to be damping, we assume that it can be written in the form

$$\varphi(u, u_t, u_x) = L_1(u, u_t, u_x)u + L_2(u, u_t, u_x)u_t + L_3(u, u_t, u_x)u_x,$$

with $L_1 \geq 0$, $L_2 \geq 0$, and $L_3 \in I\!R$. We remark that this representation of φ is not unique, and the validity of the above constraints on L_j may depend on the particular choice of representation. Many nonlinear KG equations of physical interest belong to this class. For the sine-Gordon equation with $\varphi(u) = \sin u$ and $L_1 = \frac{\sin u}{u}$ this holds true if u lies in the interval $[-\pi, \pi]$. More examples will be given in Section 4.5.

To construct a PML system for this semilinear hyperbolic system, a natural idea is to substitute the nonlinear function φ into the system (4.24)-(4.28). This treatment of nonlinear terms can be considered as a special linearization. In contrast to the direct linearization (i.e. in the equation) of nonlinear terms, it yields, however, also a nonlinear hyperbolic system in the PML zone. And for the trivial choice $\sigma \equiv 0$ one still recovers the original nonlinear KG equations on $[0, \infty)$. Therefore, we can reasonably expect that the nonlinear PML system (as a treatment of the open boundary) yields a better approximate solution than the direct linearization. This is illustrated by the numerical tests given in Section 4.5.

4.3 Two-dimensional KG equations

An advantage of PML over other absorbing boundary techniques is the easy extension to higher dimensional problems. First we consider the following nonlinear KG equation in two dimensions

$$u_{tt} = c^2(u_{xx} + u_{yy}) - \varphi(u, u_t, u_x, u_y). \tag{4.29}$$

We consider this problem on a duct with uniform rectangular tail, say $[0, \infty) \times [-b, b] \subset I\!R^2$, and derive a PML system for the exterior region $[a, \infty) \times [-b, b]$. The PML hyperbolic system for equation (4.29) can be obtained by a minor modification of the system (4.24)-(4.28):

$$u_t = w, \tag{4.30}$$

$$v_{1,t} = w_x - \sigma p, \tag{4.31}$$

$$v_{2,t} = w_y, \tag{4.32}$$

$$w_t = c^2 v_{1,x} + c^2 v_{2,y} - \varphi(u, w, v_1, v_2) - \sigma q, \tag{4.33}$$

$$p_t = w_x - (\alpha + \sigma)p, \tag{4.34}$$

$$q_t = c^2 v_{1,x} - (\alpha + \sigma)q. \tag{4.35}$$

Here we introduced $v_1 := u_x$ and $v_2 := u_y$. The boundary and initial data for the nonlinear Klein-Gordon equation (4.29) are then transformed for the new unknown functions in (4.30)-(4.35) in analogy to (4.21) and (4.22).

For two-dimensional exterior problems, the generalization is also straightforward, but in this case, more auxiliary unknown functions have to be introduced. For simplicity, we only consider the Cauchy problem. Suppose the initial functions are locally supported in a rectangular domain $[-a, a] \times [-b, b]$. We introduce two absorption functions $\sigma_x = \sigma_x(x)$ and $\sigma_y = \sigma_y(y)$. σ_x vanishes on $[-a, a]$ and σ_y vanishes on

$[-b, b]$. The modified PML system thus reads

$$u_t = w, \tag{4.36}$$
$$v_{1,t} = w_x - \sigma_x p_1, \tag{4.37}$$
$$v_{2,t} = w_y - \sigma_y p_2, \tag{4.38}$$
$$w_t = c^2 v_{1,x} + c^2 v_{2,y} - \varphi(u, w, v_1, v_2) - \sigma_x q_1 - \sigma_y q_2, \tag{4.39}$$
$$p_{1,t} = w_x - (\alpha + \sigma_x) p_1, \tag{4.40}$$
$$p_{2,t} = w_y - (\alpha + \sigma_y) p_2, \tag{4.41}$$
$$q_{1,t} = c^2 v_{1,x} - (\alpha + \sigma_x) q_1, \tag{4.42}$$
$$q_{2,t} = c^2 v_{2,y} - (\alpha + \sigma_y) q_2. \tag{4.43}$$

Initial boundary value problems can be dealt with analogously.

4.4 Numerical scheme

In the real implementation, the PML zone has to be truncated at some $x = \tilde{a} > a$, and an auxiliary boundary condition is needed on the exterior PML boundary. Owing to the strong damping property of the PML zone, a homogeneous Dirichlet or Neumann boundary condition can serve this purpose, provided it yields a well-posed truncated problem. For our problems we use a homogeneous boundary condition for the one incoming the characteristic of the hyperbolic part of the PML system. In the example (4.24)-(4.28) this reads

$$cv(\tilde{a}, t) + w(\tilde{a}, t) = 0. \tag{4.44}$$

To design a suitable numerical scheme, we confine ourselves to the one-dimensional PML system (4.24)-(4.28), with φ possibly nonlinear. This is a semilinear, weakly hyperbolic system with a stiff damping term when σ is large. To guarantee stability, a very small time step must be used for any explicit numerical scheme, which would greatly increase the computational cost. In this paper, we employ the well-developed implicit-explicit (IMEX) Runge-Kutta schemes to overcome this difficulty.
Let us first consider a large ODE system

$$U_t = H(U) + S(U), \tag{4.45}$$

which might come from the discretization of the spatial derivatives of a complex PDE problem. H corresponds to the hyperbolic term, while S (which is supposed

to be stiff) corresponds to the diffusion or –as in our case– reaction terms. This means that there are two significantly different time scales involved in (4.45). A normal implicit ODE solver can guarantee numerical stability. However, at each time step, a nonlinear algebraic system with a large number of unknowns (most of them coupled) have to be solved which makes these schemes inapplicable. The idea of IMEX Runge-Kutta is to use an explicit scheme on H and an implicit scheme on S. The overall scheme is still implicit. But in this case the resulting nonlinear algebraic system can be decoupled into many smaller nonlinear algebraic systems. Thus, the overall computational cost is acceptable.

Suppose that Δt is the time step and the numerical solution at the n-th step is U_n. The IMEX Runge-Kutta approximation for the $n+1$-st step is realized by

$$U^{(i)} = U_n + \Delta t \sum_{j=1}^{i-1} \tilde{a}_{ij} H(U^{(j)}) + \Delta t \sum_{j=1}^{\nu} a_{ij} S(U^{(j)}); \quad i = 1, \cdots, \nu, \tag{4.46}$$

$$U_{n+1} = U_n + \Delta t \sum_{i=1}^{\nu} \tilde{w}_i H(U^{(i)}) + \Delta t \sum_{i=1}^{\nu} w_i S(U^{(i)}). \tag{4.47}$$

Here, the matrices $\tilde{A} = (\tilde{a}_{ij})$, $\tilde{a}_{ij} = 0$ for $j \geq i$ and $A = (a_{ij})$ are $\nu \times \nu$ matrices. $\tilde{A}$ and $\tilde{W} = (\tilde{w}_i)$ correspond to an explicit Runge-Kutta scheme, while A and $W = (w_i)$ correspond to an implicit Runge-Kutta scheme. For efficiency of solving the nonlinear algebraic system corresponding to the implicit part, it is natural to consider diagonally implicit Runge-Kutta (DIRK) schemes for the damping term, i.e., $a_{ij} = 0$ for $j > i$.

For a second-order time integration, we use the following implicit-explicit midpoint scheme

$$U^{(1)} = U_n + \frac{\Delta t}{2} \left(H(U_n) + S(U^{(1)}) \right), \tag{4.48}$$

$$U_{n+1} = U_n + \Delta t (H(U^{(1)}) + S(U^{(1)})). \tag{4.49}$$

The reader is referred to [**9, 68**] for more details on IMEX Runge-Kutta schemes.

The PML system (4.24)-(4.28) for nonlinear KG equations can be written as

$$Q_t = AQ_x + S(Q, x), \tag{4.50}$$

with $Q = (u, v, w, p, q)^T$, and

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & c^2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & c^2 & 0 & 0 & 0 \end{pmatrix}, \; S = \begin{pmatrix} w \\ -\sigma(x)p \\ -\varphi(u, w, v) - \sigma(x)q \\ -(\alpha(x) + \sigma(x))p \\ -(\alpha(x) + \sigma(x))q \end{pmatrix}.$$

With mesh size Δx, the spatial variable is discretized using a second order finite-volume scheme. We introduce the grid points as

$$x_j = j\Delta x, \; x_{j+\frac{1}{2}} = x_j + \frac{1}{2}\Delta x, \quad j \in I\!N_0$$

and use the standard notations

$$Q_j = Q(x_j, t), \; \bar{Q}_j = \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} Q(x, t)dx.$$

By integrating (4.50) over $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ and dividing by Δx we obtain

$$\bar{Q}_{j,t} = \frac{1}{\Delta x}\big(AQ(x_{j+\frac{1}{2}}, t) - AQ(x_{j-\frac{1}{2}}, t)\big) + \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} S(Q(x, t), x)dx.$$

Within second order accuracy, the average of the damping term $S(Q, x)$ can be taken as

$$\frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} S(Q(x, t), x)dx \approx \begin{pmatrix} \bar{w}_j \\ -\sigma_j\bar{p}_j \\ -\varphi(\bar{u}_j, \bar{w}_j, \bar{v}_j) - \sigma_j\bar{q}_j \\ -(\alpha_j + \sigma_j)\bar{p}_j \\ -(\alpha_j + \sigma_j)\bar{q}_j \end{pmatrix} = S(\bar{Q}_j, x_j).$$

The treatment of the hyperbolic part is rather standard. The numerical flux at the interface of two grid cells is obtained by solving a Riemann problem. The left and right values at each interface point is obtained by a linear reconstruction from the average cell values. Since we are studying nonsmooth solutions, a suitable limiter such as the minmod limiter should be used.

Finally, the resulting ODE system is then solved by (4.48)-(4.49) for the time integration. Higher dimensional problems can be solved analogously.

4.5 Numerical examples

4.5.1 Example 1: one-dimensional linear KG equation

We consider

$$u_{tt} = u_{xx} - u, \quad x > 0, \quad t > 0, \tag{4.51}$$

with

$$u_0 = u_1 \equiv 0, \; u(0,t) = \sin(\omega t), \qquad t > 0, \tag{4.52}$$

for some fixed $\omega > 0$. The solution of (4.51)-(4.52) corresponds to an evanescent wave if $\omega < 1$ and to a traveling wave if $\omega > 1$. The computational domain is set to $[0,1]$. The step sizes are chosen small enough ($\Delta t = \Delta x = 0.0025$) such that the discretization errors are negligible compared to the errors due to the ABC. The thickness of the PML zone is 0.2 and the absorption function is chosen as

$$\sigma(x) = \begin{cases} 0, & 0 \le x \le 1, \\ 10^5 (x-1)^3, & x > 1. \end{cases}$$

Now we shall compare the exact solution of (4.51) to the solution of the modified PML system (4.16)-(4.20) with truncation at $x = 1.2$, i.e. with the characteristic boundary condition (4.44). For this comparison, we take as "exact" reference solution the numerical solution of (4.51) computed on the large spatial domain $[0,10]$ with the boundary condition $u(10) = 0$. Due to the wave speed 1, boundary effects will only affect $u|_{[0,1]}$ for $t \ge 19$. Figure 4.2 compares the $L^2(0,1)$-error, for $0 \le t \le 10$, for different choices of the phase shift α. It shows that using a positive phase shift

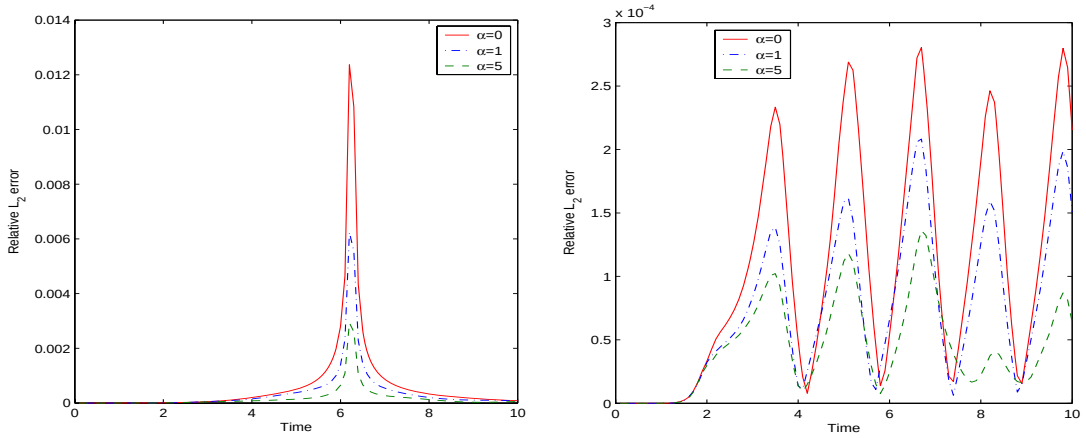


Figure 4.2: The influence of phase shift. Left: $\omega = 0.5$. Right: $\omega = 2$.

leads to a more accurate numerical solution. Besides, if we compare the right and left plots, we notice that the PML technique yields better results for the traveling wave ($\omega = 2$) than for the evanescent wave (notice the different scales!).
In Figure 4.3 we depict the approximate solutions of (4.51)-(4.52) with $\omega = 2$ for the two cases $\alpha = 0$ and $\alpha = 1$, at different points in time. Focusing on the domain

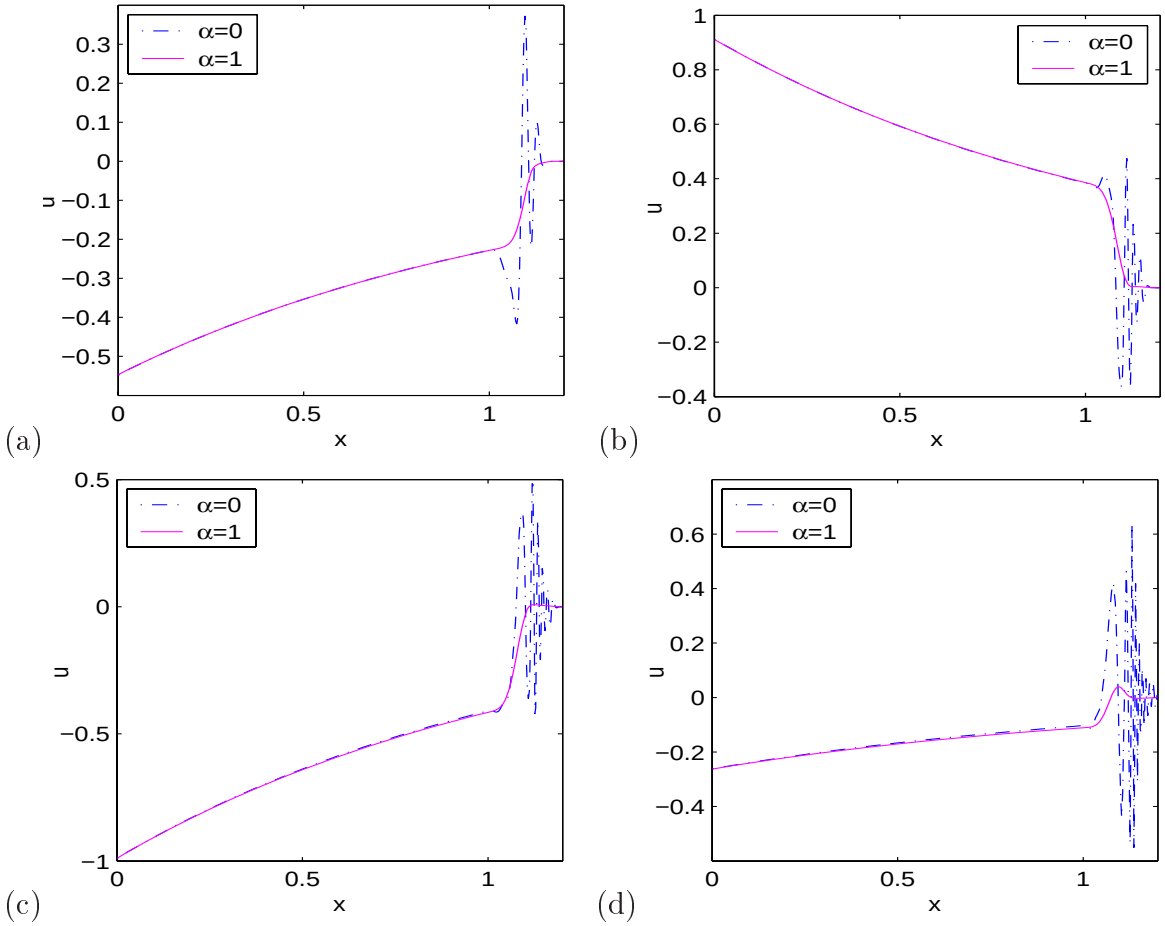


Figure 4.3: Numerical solutions for $\omega = 2$ with different phase shifts: (a) $t = 20$; (b) $t = 40$; (c) $t = 60$; (d) $t = 100$.

$[0, 1]$, we notice that the solutions in the two cases match very well. Even at $t = 100$, only a minor difference can be observed. However, on the PML zone, $[1, 1.2]$, the solution with phase shift is much smoother. This effect of the phase shift obviously stabilizes the solution with respect to the truncation of the PML zone.

4.5.2 Example 2: one-dimensional sine-Gordon equation

As discussed in Section 2, the presented PML derivation can be considered as some kind of linearization for the governing equation. This poses the following question: How is the performance of the PML linearization compared to the direct linearization of nonlinear terms in the equation? To this end we consider the one-dimensional sine-Gordon equation

$$u_{tt} = u_{xx} - \sin u, \quad x > 0,\ t > 0,$$

with initial and boundary conditions (4.52). A direct linearization of the nonlinear damping term about $u \equiv 0$ yields the same governing equation (4.51) of Example 1. The evolution of $L^2(0,1)$-error in Figure 4.4 shows that the PML linearization

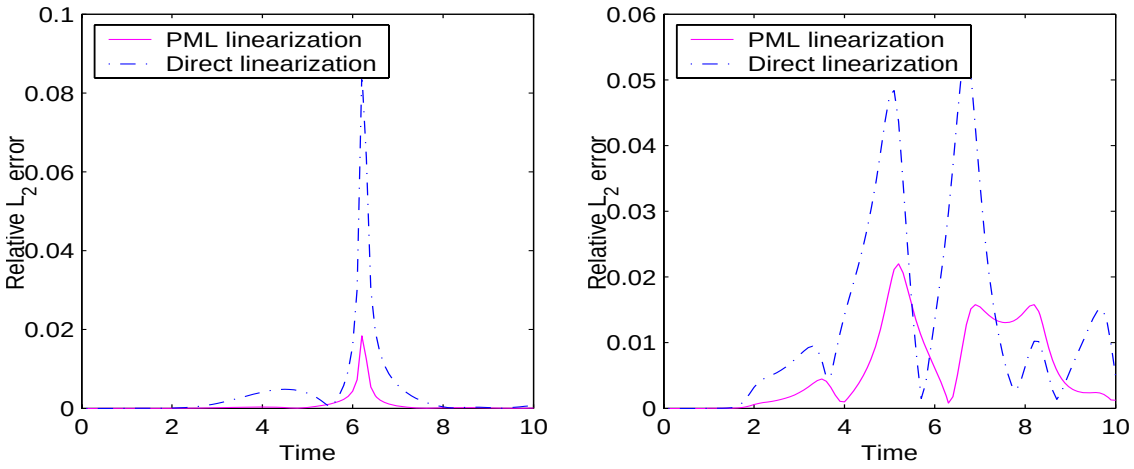


Figure 4.4: Comparison between PML linearization and direct linearization. Left: $\omega = 0.5$. Right: $\omega = 2$.

performs better than the direct linearization.

4.5.3 Example 3: comparison of PML with absorbing boundary conditions

In this test, we compare the PML linearization technique with the local absorbing boundary conditions proposed by Szeftel [**73**]. Consider the one-dimensional nonlinear KG equation

$$u_{tt} = u_{xx} - \varphi(u, u_t, u_x).$$

The initial functions

$$u_0(x) = x^3(2-x)^3\chi_{[0,2]},\ u_1(x) = 3x^2(2-x)^2(x-1)\chi_{[0,2]}, \tag{4.53}$$

are compactly supported in $[0,2]$. Four different damping terms will be considered

Case A: $\varphi = u_t$;

Case B: $\varphi = u^3$;

Case C: $\varphi = u + u^3$;

Case D: $\varphi = u^2 u_t$.

The computational parameters were chosen small enough ($\Delta x = 0.001$ and $\Delta t = 0.0005$) such that we can ignore the discretization error. The phase shift α is set to 1. The "exact solution" is taken as the numerical solution obtained in an enlarged domain $[-10, 12]$ with the same computational parameters. Its restriction to $[0,2]$ are plotted in Figure 4.5. To measure the accuracy of these "exact" reference solutions, we compared them to solutions obtained with $2\Delta x$ and $2\Delta t$: This yields an absolute L^∞-error of the order of 10^{-8} (with $u_0(1) = 1$). Thus, the computed reference solutions are sufficiently accurate.
In this Example, we define the L^2-error function as in Szeftel [**73**] by

$$\frac{||u(t,\cdot) - u_{ex}(t,\cdot)||_{L^2}}{||u_0||_{L^2} + ||u_1||_{L^2}}.$$

In our computation, we set the thickness of PML zone as 0.1 at both ends. The absorption function is set to be

$$\sigma(x) = \begin{cases} -10^6 x^3, & x < 0, \\ 0, & 0 < x < 2, \\ 10^6(x-2)^3, & x > 2. \end{cases}$$

Figure 4.6 compares the $L^2(0,2)$-error of the ABCs derived in [**73**] based on pseudodifferential approach and that of the PML linearization technique for the Case A. It is clear that the PML linearization presents much more accurate numerical solutions than others.This is expected, since for linear problems the solution to the modified PML system without truncating the PML zone is exactly the same as the exact solutions in the continuous level.

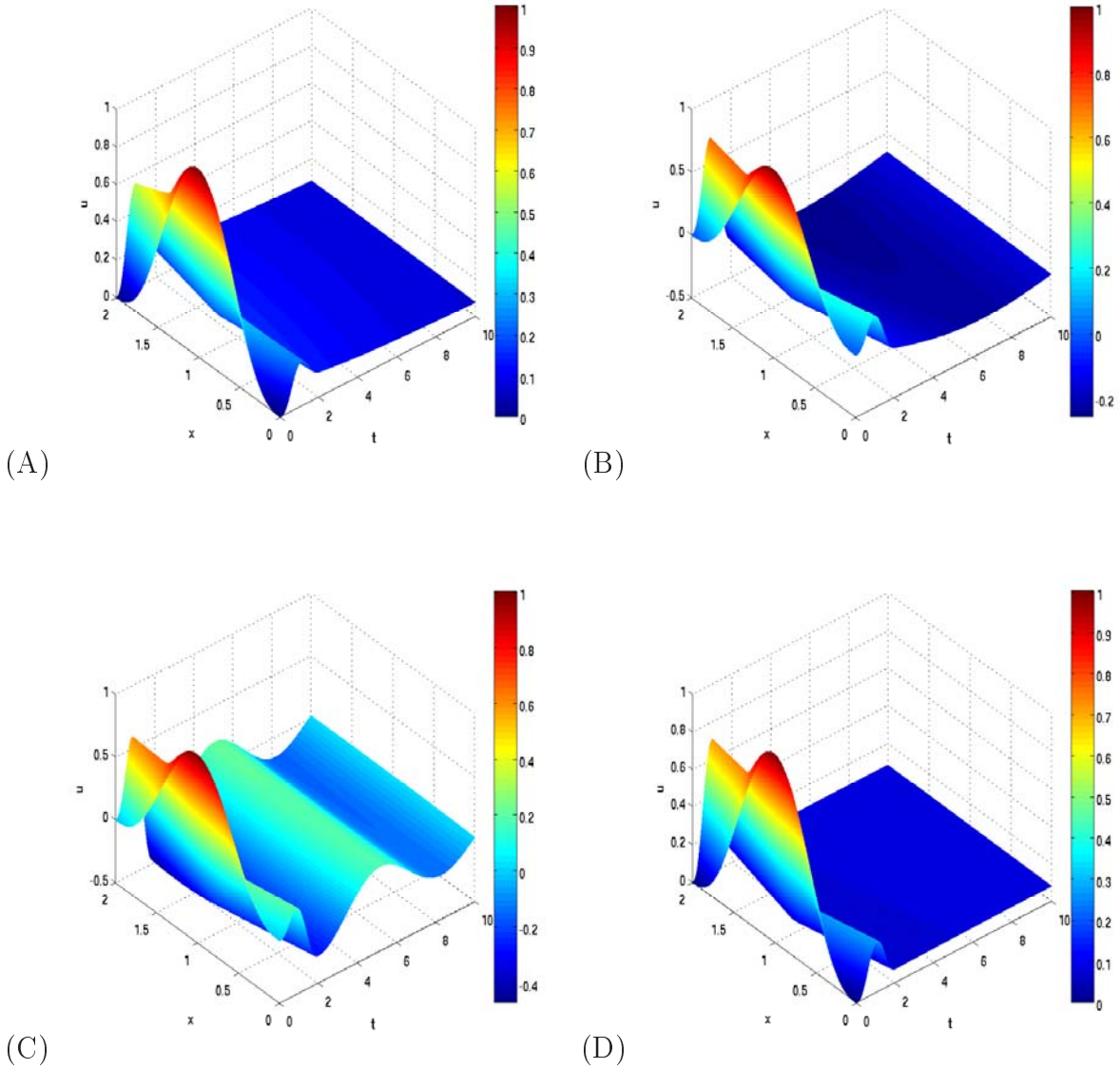


Figure 4.5: Exact solutions on $[0, 2]$ for the Cases A-D.

For Cases B and C, the ABCs derived by paradifferential and pseudodifferential calculus are the same **[73]**. Figure 4.7 shows the comparison between the PML linearizations, direct linearization, the first and the second-order ABCs. In Figure 4.7-B, we notice that the PML linearization presents a better approximations over both the direct linearization and the first-order ABC, which are identical in this case, and a competitive one with the second-order ABCs. Figure 4.7-C shows that the PML linearization presents the best approximations with relative error less than

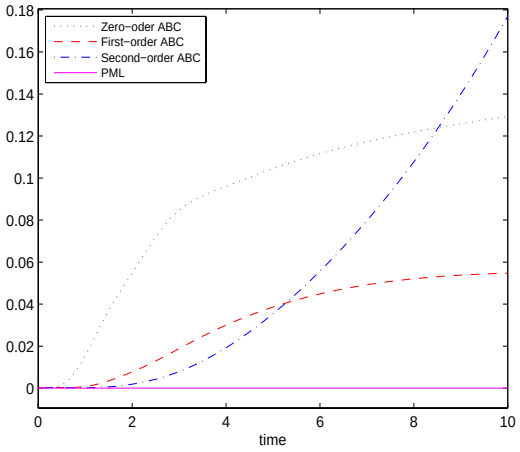


Figure 4.6: Comparison of the $L^2(0,2)$-error between different ABCs for Case A.

2%.

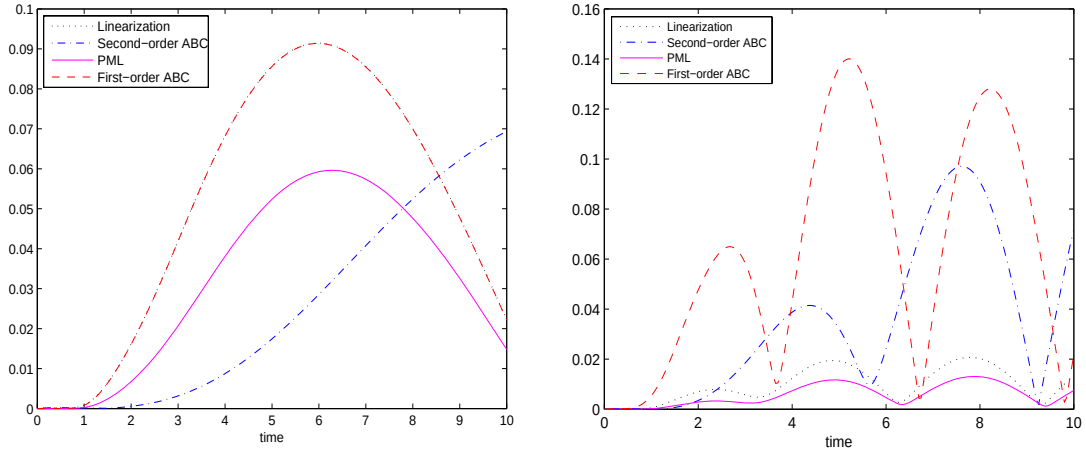


Figure 4.7: Comparison of the $L^2(0,2)$-error between different ABCs. for the Cases B (left) and C (right).

The Case D is shown in Figure 4.8 in which a competitive approximations can be observed between the PML linearization and the first-order ABC based on the paradifferential approach. The best performance is that of the second-order paradifferential ABC. However, Figure 4.10-D shows the advantage of PML over long time computations. The exact solutions for long time is taken as the numerical solutions obtained in a large domain $[-55, 57]$, with the same step sizes as that in the short time computations. The exact solutions of the Cases B,C, and D restricted to $[0, 2]$ are plotted in Figure 4.10-Right. The left part of Figure 4.10 shows the comparison

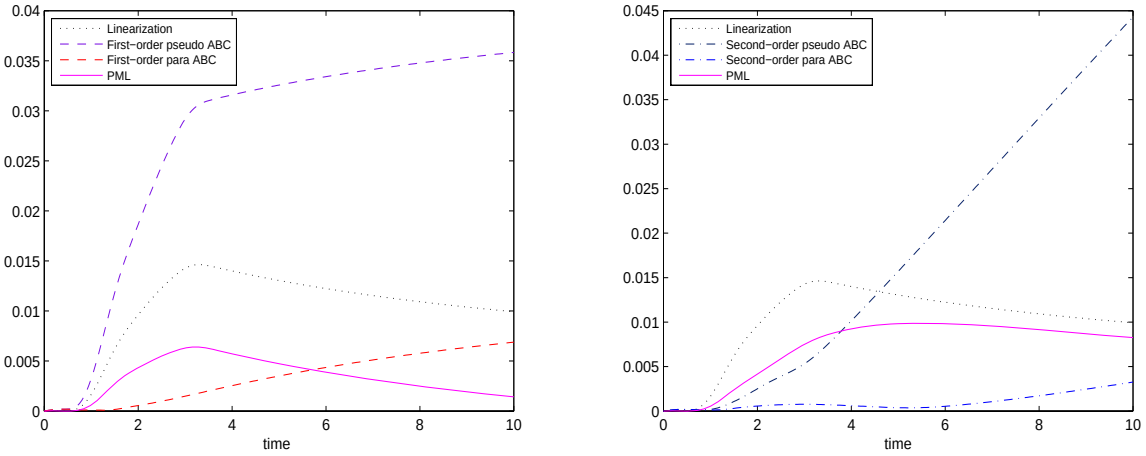


Figure 4.8: Comparison of the $L^2(0,2)$-error between different ABCs for Case D.

of the $L^2(0,2)$-error over long time.

Another remarkable advantage of our PML linearization over the local ABCs lies in the easy extension to the high-dimensional problems, as the next example shows.

4.5.4 Example 4: two-dimensional exterior problem

Consider the two-dimensional nonlinear KG equation

$$u_{tt} = u_{xx} + u_{yy} - u^3,$$

with initial condition

$$u_0(x,y,0) = 0,\ u_1(x,y,0) = \begin{cases} 10, & x^2+y^2 < 3, \\ 0, & \text{otherwise.} \end{cases}$$

The physically interested domain is $[-2,2]\times[-2,2]$ and the computational domain including the PML zone is set to be $[-2.8,2.8]\times[-2.8,2.8]$. The same absorption function as that in Example three is used in both x and y directions. The computational parameters are set to be $\Delta x = \Delta y = 0.014$ and $\Delta t = 0.007$. For comparison, we take the exact solution as the numerical solution by solving the above problem in an enlarged domain $[-5.6,5.6]$ with the same computational parameters.

Figure 4.9 compares the relative $L^2((-2,2)^2)$-errors due to both PML linearization and direct linearization. Again, the PML linearization shows an advantage over the direct linearization.

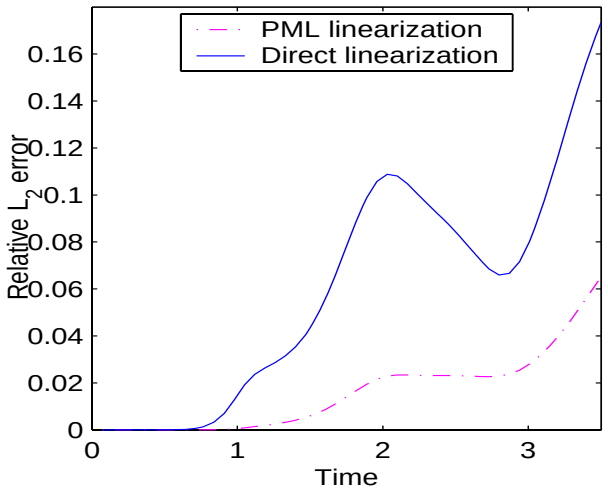


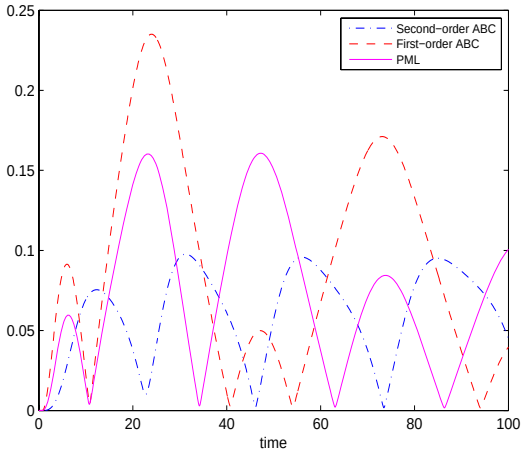
Figure 4.9: Comparison of relative errors from PML linearization and direct linearization.

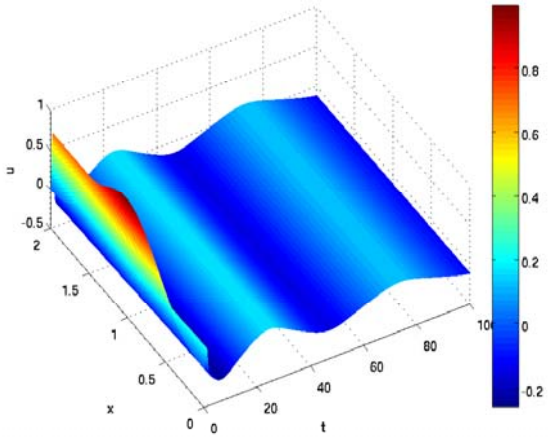
In Figure 4.11, we draw the contour plots for the exact solution, the approximate solution with PML linearization, and that with direct linearization. The difference between the exact solution and the numerical solution with PML linearization can hardly be detected, while in the case of direct linearization this difference is obvious.

4.6 Conclusion

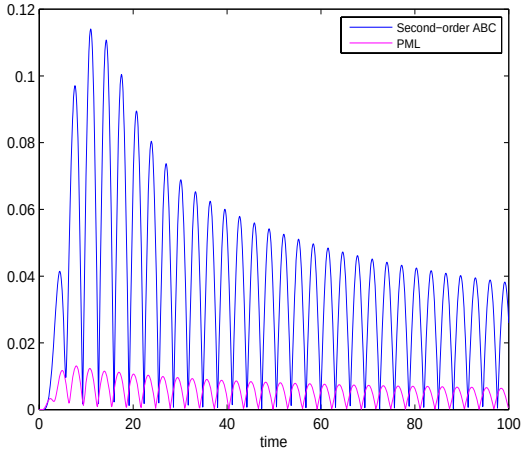
In this paper, we have proposed a PML linearization technique to numerically handle nonlinear Klein-Gordon equations. We explained in detail the design of the modified PML system and illustrated the importance of the phase shift parameter in numerical tests. The PML linearization can be considered as a special kind of linearization technique for nonlinear problems. Compared with the direct linearization, numerical tests in one and two dimensions for various nonlinear damping terms have shown its advantage. Our numerical tests have also shown the efficiency of the PML linearization over other local ABCs obtained by paradifferential and pseudodifferential approaches. For most examples the PML method gives better results. Another advantageous property of this technique lies in its easy generalization to higher dimensions.

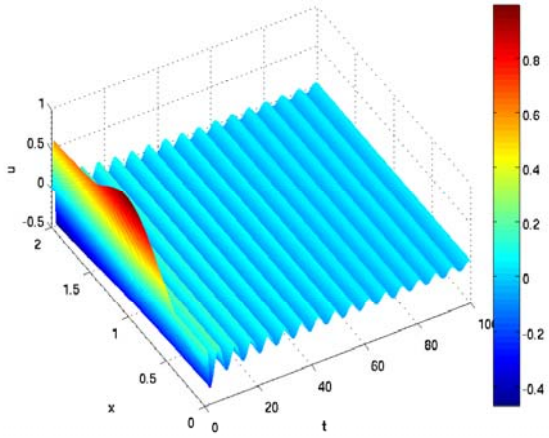
(B)



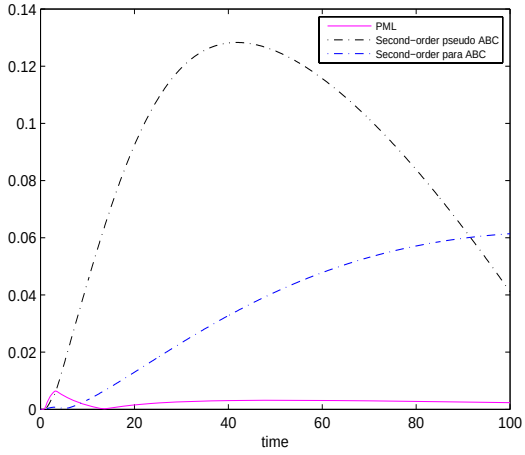


(C)





(D)



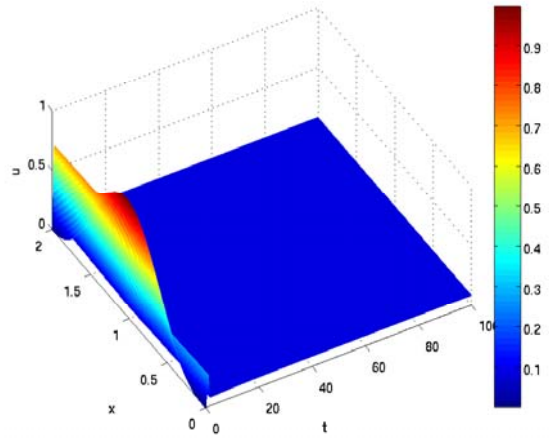


Figure 4.10: Left: Comparison of the $L^2(0,2)$-error between different ABCs. Right: Exact solution.

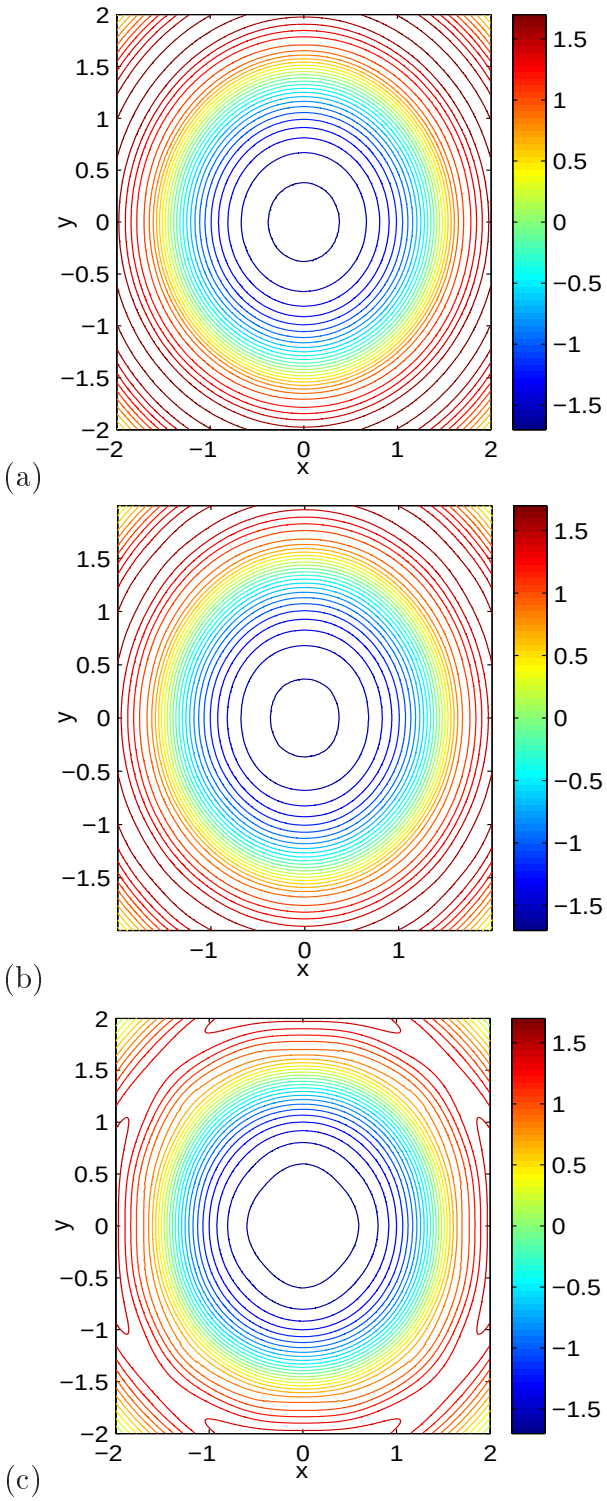


Figure 4.11: Contour plots at $t = 3.5$. (a) Exact solution; (b) PML linearization; (c) Direct linearization.

Chapter 5

Absorbing PML boundary layers for the nonlinear Euler equations

Abstract

In this paper a PML absorbing boundary condition (ABC) is presented for the nonlinear Euler equations. There are two steps involved. First, the PML technique is applied to the Euler equations linearized about uniform and parallel flows. Then the nonlinear PML equations are formed by replacing the linearized flux functions with their nonlinear counterparts. Since a stiff source term gets involved in the nonlinear PML equations, an Implicit-Explicit Runge-Kutta scheme is recommended to compute numerical solutions. Some tests are performed, and the results demonstrate the effectiveness of the proposed PML ABC.

5.1 Introduction

[1]Wave propagation in unbounded domain appears in many fields of application, such as aeroacoustics, electromagnetics, and seismics [**29**]. A main aim of numerical simulation for these kind of problems is to resolve the wave field on a small part of domain which bears some special physical interest. Therefore it is a natural practice to limit the computational domain by introducing some artificial boundaries, which

[1]The content of this chapter is a joint work with C. Zheng (cf. [**83**])

necessitates suitable boundary conditions to be imposed. A simple choice of these boundary conditions, for example, the radiation condition at infinity, works well if the computational domain is made large enough such that the simulation terminates before the wave reflection from the artificial boundaries comes into effect. But this treatment generally leads to a heavy computation burden in terms of time and storage. Alternatively, the physical domain can be tailored more closely, and thus more delicate boundary conditions must be designed.

We suppose that the solution near the artificial boundary consists only of outgoing waves. In this case, a good boundary condition should not only present a wellposed problem, but also mimic the absorption of waves originating from the interior, and prevent the energy from being reflected by the artificial boundary. Right in this context, such a boundary condition is usually called absorbing boundary condition (ABC). Other names of same spirit are also popularly used in the literature, such as nonreflecting, transparent and open boundary conditions. Besides, from the computational aspect, a good ABC should also be inexpensive to implement.

Different ABCs have been developed over the last three decades. The readers are referred to [**29, 40, 78, 41, 18**] for detailed review. In general the ABCs can be classified into two categories: PDE-based and material-based. PDE-based ABCs are imposed exactly on prescribed artificial boundaries, and obtained either by factorizing the field equation and allowing only the outgoing waves, or by solving the exterior wave problems in an analytical or semi-analytical way. Material-based ABCs employ a different philosophy. Instead of solely using artificial boundaries to limit the computational domain, they turn to a finite-thickness lossy material to annihilate the outgoing waves. These kind of ABCs have been utterly renovated since 1994 when Bérenger [**13**] proposed the PML for computational electromagnetics. Soon after, Chew and Weedon [**15**] presented an elegant explanation that the PML essentially equals to some coordinate stretching from the real axis to the complex plane, making all waves damped in the PML absorbing layers under the new coordinate system. Abarbanel and Gottlieb revealed in [**1**] that the original PML formulation in [**13**], which is based on the split physical variables, is only weakly well-posed.

Of particular interest in this work is the application of the PML for the Euler equations. Hu [**51**] was the first to consider such a problem. Based on the idea in [**13**], he proposed a PML formulation for the Euler equations linearized about uniform flow. Goodrich and Hagstrom [**34**] presented a different formulation, and reported that in some special cases their formulation admits unstable growing wave

modes, which implies the existence of some inherent instability. This was confirmed both theoretically and numerically by many authors, including Hesthaven [**49**] and Tam *et al.* [**75**]. Considering this point, in 1999, Abarbanel *et al.* [**3**] first derived a well-posed PML by using the unsplit physical variables. Hu [**52**] presented another stable formulation by performing a space-time transformation before applying the PML technique. Later, he extended this idea to the parallel shear flow [**54**] and proposed a numerical procedure to find such a transformation. Recently, Hu [**55**] and his collaborators [**57**] even considered the nonlinear Euler equations and the Navier-Stokes equations. Hagstrom [**42**] constructed a general PML formulation for linear hyperbolic systems, and with Nazarov [**43, 44**] he considered the parallel flow and studied the dependence of growth rate within the PML layer on the involved parameters.

In this paper we study the PML technique for the nonlinear Euler equations. We consider both the ducted flow problems for which only vertical x-layers to the left and right of the physical domain are needed, and the open flow problems for which horizontal y-layers are further needed at the top and bottom boundaries, see Figure 5.1. A new PML formulation is presented, which has an advantage that the numerical schemes for nonlinear conservation laws can be adapted easily. We propose a strategy for choosing the parameters involved in this formulation and describe an Implicit-Explicit Runge-Kutta scheme with which we solve the presented numerical examples.

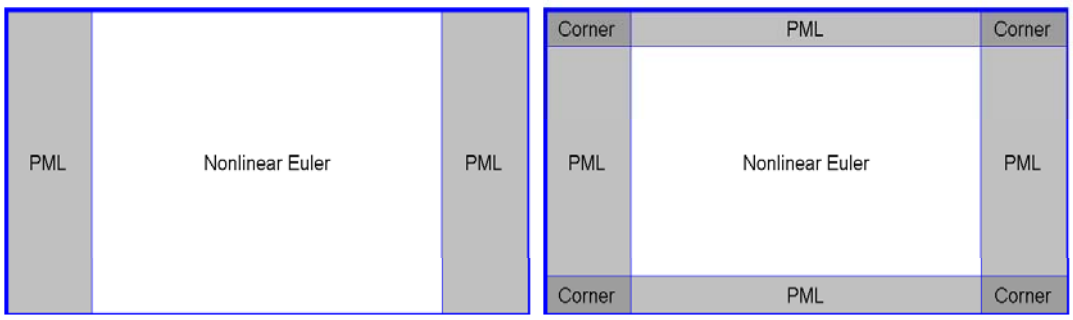


Figure 5.1: Two schematics of unbounded domain problems. Left: ducted flow. Right: open flow.

The organization of this paper is as follows. In Section 2, we present PML formulations for the linearized Euler equations with both uniform and nonuniform but parallel flows. In Section 3, we discuss how to construct the final PML formulation for the nonlinear Euler equations. In Section 4, we explore one strategy to solve the

nonlinear PML equations. Numerical examples are presented in Section 5, and we conclude this paper in Section 6.

5.2 PML formulations for the linearized Euler equations

The two-dimensional nonlinear Euler equations describing the physical laws of conservation of mass, momentum and energy, read

$$q_t + f(q)_x + g(q)_y = 0, \tag{5.1}$$

where

$$q = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ E \end{pmatrix}, \ f(q) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ (E+p)u \end{pmatrix}, \ g(q) = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ (E+p)v \end{pmatrix}. \tag{5.2}$$

The unknown vector $q(t, x, y)$ from $\mathbf{R}^2 \times [0, +\infty)$ into Ω describes the state of the gas as a function of time and space. The set Ω is called the set of states and f and g are the convective fluxes in the x- and y-directions, respectively. Here ρ, u, v, E and p denote density, velocity components in the x- and y-directions, total energy and pressure. For a perfect gas

$$p = (\gamma - 1)\left(E - \frac{\rho}{2}(u^2 + v^2)\right), \tag{5.3}$$

where γ is the ratio of specific heats.

When the flow near the boundary can be considered as a constant mean flow plus a small amplitude fluctuation, it is justifiable to linearize the nonlinear Euler equations (5.1) and approximate the fluctuation part by the solution to the resulting linearized Euler equations. A typical difficulty in applying the PML for the parallel flow when compared with the uniform flow is to damp all those waves which travel in the PML absorbing layers. We will study this issue in the following.

5.2.1 Uniform flow

Linearizing the Euler equations with a constant state $q_0 = (\rho_0, \rho_0 u_0, \rho_0 v_0, E_0)$ yields the following hyperbolic system with constant coefficients

$$\tilde{q}_t + A\tilde{q}_x + B\tilde{q}_y = 0, \tag{5.4}$$

where

$$\tilde{q} = q - q_0,\ A = f'(q_0),\ B = g'(q_0) \tag{5.5}$$

and

$$f'(q) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -u^2 + \frac{1}{2}(\gamma-1)(u^2+v^2) & (3-\gamma)u & -(\gamma-1)v & \gamma-1 \\ -uv & v & u & 0 \\ u\left[\frac{1}{2}(\gamma-1)(u^2+v^2) - H\right] & H-(\gamma-1)u^2 & -(\gamma-1)uv & \gamma u \end{pmatrix},$$

$$g'(q) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -uv & v & u & 0 \\ -v^2 + \frac{1}{2}(\gamma-1)(u^2+v^2) & -(\gamma-1)u & (3-\gamma)v & \gamma-1 \\ v\left[\frac{1}{2}(\gamma-1)(u^2+v^2) - H\right] & -(\gamma-1)uv & H-(\gamma-1)v^2 & \gamma v \end{pmatrix}.$$

Here, $H = (E+p)/\rho$ is the enthalpy. We confine to the subsonic flow, i.e., $u_0^2 + v_0^2 < c_0^2 = \frac{\gamma p_0}{\rho_0}$. For the linear system (5.4), the standard modal analysis can be applied. Suppose the modal solution behaves like

$$\tilde{q} = Qe^{st+\lambda x+iky}. \tag{5.6}$$

Here s with $\Re(s) > 0$ is the variable in the Laplace-transformed space, λ is the complex wave number in the x-direction, and k corresponds to the real wave number in the y-direction. Q is a vector depending on s, λ and k. Substituting (5.6) into equation (5.4) we get

$$(sI + \lambda A + ikB)\tilde{q} = 0. \tag{5.7}$$

Equation (5.7) means that for any fixed s and k, λ is a generalized eigenvalue of equation (5.7), and $\tilde{q}$ is a corresponding eigenvector. A nontrivial solution can be obtained if λ satisfies the following dispersive relation

$$\det(sI + \lambda A + ikB) = 0, \tag{5.8}$$

which has three solutions

$$\lambda_1 = -\frac{s + ikv_0}{u_0}, \tag{5.9}$$

$$\lambda_- = \frac{(s+ikv_0)u_0 - c_0\sqrt{(s+ikv_0)^2 + k^2(c_0^2-u_0^2)}}{c_0^2 - u_0^2}, \tag{5.10}$$

$$\lambda_+ = \frac{(s+ikv_0)u_0 + c_0\sqrt{(s+ikv_0)^2 + k^2(c_0^2-u_0^2)}}{c_0^2 - u_0^2}. \tag{5.11}$$

Here, λ_- and λ_+ correspond to the acoustic waves and λ_1 to the entropy and the vorticity waves simultaneously. The direction of these waves depends on the sign of the real part of their corresponding λ: left-going if positive and right-going if negative. Thus, the acoustic waves corresponding to λ_- and λ_+ always move to the right and the left separately. If $u_0 > 0$, the entropy and vorticity waves are right-going, and if $u_0 < 0$, they are left-going. For convenience, from now on we will refer to "the sign of the real part of λ" simply as "the sign of λ".

The essence of the PML lies in constructing new wave equations to which the modal solutions not only are continuations of those to the original wave equations, but also decay exponentially faster in their propagating directions at the same time. This is usually accomplished with the following coordinate transformation

$$x \longrightarrow x' = x + \frac{f}{\lambda}\int_{x_0}^{x} \sigma(z)dz, \ x \geq x_0,$$

where $\sigma \geq 0$ is called the absorption coefficient, and f is a mapping from λ to the complex plane. Under this transformation, the modal solution (5.6) is changed to

$$\tilde{q} = Qe^{st+\lambda x'+iky} = Qe^{st+\lambda x+iky}e^{f\int_{x_0}^{x}\sigma(z)dz}. \tag{5.12}$$

Obviously, the point to make this modified modal solution damp faster than (5.6) in its traveling direction, is to ensure f has the same sign of λ. Following the idea of Hagstrom [**42**], we set

$$f = \left(\frac{\lambda}{s+ikv_0} - \mu\right)\frac{s+ikv_0}{s+ikv_0+\alpha}, \tag{5.13}$$

where $\mu := \frac{u_0}{c_0^2-u_0^2}$, and $\alpha \geq 0$ is called the phase shift parameter. f indeed has the same sign of λ, as can be verified directly.

From (5.12) and (5.13) , we have

$$\tilde{q}_x = \left(\lambda + \frac{\lambda\sigma}{s+ikv_0+\alpha} - \sigma\mu\frac{s+ikv_0}{s+ikv_0+\alpha}\right)\tilde{q}.$$

Thus then,

$$\lambda\tilde{q} = \frac{s+ikv_0+\alpha}{s+ikv_0+\alpha+\sigma}\left(\tilde{q}_x + \sigma\mu\frac{s+ikv_0}{s+ikv_0+\alpha}\tilde{q}\right).$$

Substituting the above expression of $\lambda\tilde{q}$ into (5.7) we get

$$s\tilde{q} + \frac{s+ikv_0+\alpha}{s+ikv_0+\alpha+\sigma}A\left(\tilde{q}_x + \sigma\mu\frac{s+ikv_0}{s+ikv_0+\alpha}\tilde{q}\right) + ikB\tilde{q} = 0.$$

Introducing an auxiliary vector function

$$w = -\frac{1}{s + ikv_0 + \alpha + \sigma} A\left(\tilde{q}_x + (\alpha + \sigma)\mu\tilde{q}\right),$$

and going back to the physical space, we derive the PML equations for the linearized Euler equations (5.4)

$$\tilde{q}_t + A\tilde{q}_x + B\tilde{q}_y + \sigma\mu A\tilde{q} + \sigma w = 0, \tag{5.14}$$

$$w_t + A\tilde{q}_x + v_0 w_y + (\alpha + \sigma)\mu A\tilde{q} + (\alpha + \sigma)w = 0. \tag{5.15}$$

It has been proved in Appelö *et al.*[**7**] that the PML equations (5.14)-(5.15) are well-posed. For constant σ, Appelö *et al.* [**7**] showed that the specific choice of $\mu = \frac{u_0}{c_0^2 - u_0^2}$ is necessary and sufficient to guarantee the asymptotic stability of (5.14)-(5.15). This issue was also studied in Hagstrom [**42**] and Hu [**52**].

5.2.2 Parallel flow

In many cases, linearization with a nonuniform parallel flow $q_0 = (\rho_0, \rho_0 u_0, 0, E_0)$ is more reasonable. The resulting linearized Euler equations read

$$\tilde{q}_t + A\tilde{q}_x + B\tilde{q}_y + B^*\tilde{q} = 0, \tag{5.16}$$

where $\tilde{q}$, A and B are given as in (5.5). The matrix B^* originates from the nonuniformity of the flow in the y-direction. We do not intend to give the explicit expression of B^*, since it is never used in the mathematical analysis and the numerical computation.

To analyze (5.16), it is a common practice to reformulate it with the primitive variables as

$$r_t + Cr_x + Dr_y + Er = 0, \tag{5.17}$$

with $r = (\tilde{\rho}, \tilde{u}, \tilde{v}, \tilde{p})^T = (\rho - \rho_0, u - u_0, v, p - p_0)^T$ and

$$C = \begin{pmatrix} u_0 & \rho_0 & 0 & 0 \\ 0 & u_0 & 0 & \frac{1}{\rho_0} \\ 0 & 0 & u_0 & 0 \\ 0 & \gamma p_0 & 0 & u_0 \end{pmatrix}, \; D = \begin{pmatrix} 0 & 0 & \rho_0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\rho_0} \\ 0 & 0 & \gamma p_0 & 0 \end{pmatrix}, \; E = \begin{pmatrix} 0 & 0 & \rho_0' & 0 \\ 0 & 0 & u_0' & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.18}$$

Unlike the linear equations (5.4) with constant coefficients, the modal analysis of (5.17) can be, in general, only performed numerically. Since the numerical computation relies strongly on the physical boundary condition in the y-direction, we have

to consider this issue case by case.

We confine ourselves to the y-periodic flow with a nonuniform velocity $(u_0, 0)$. The period is L. ρ_0 and p_0 are assumed constant. In this case, a simple normalization can be made to give $\rho_0 = 1$ and $p_0 = \frac{1}{\gamma}$. We presume that this normalization has already been done.

Let the modal solution of (5.17) be

$$r = e^{st+\lambda x}\phi, \tag{5.19}$$

where ϕ is a vector function depending on s, λ and the y-coordinate. Substituting (5.19) into (5.17) we get

$$(sI + \lambda C + E + D\partial_y)\phi = 0, \tag{5.20}$$

where I is the 4×4 identity matrix. For any fixed s with $\Re s > 0$, there are an infinite number of generalized eigenvalues λ and corresponding eigenfunctions ϕ satisfying the above equation. We compute those λ whose eigenfunctions can be well-approximated with less than 121 Fourier modes.

To understand how to extend the PML technique to the nonuniform flow, we perform first the numerical modal analysis for a uniform mean state with $u_0 = 0.5$, even though it has been analyzed in the last subsection. Figure 5.2 shows the results for different values of s (we set $L = 2\pi$). Let us look at Figure 5.2-b. The first subplot shows all admissible generalized eigenvalues λ. For a clear distinction, we have marked the negative and positive eigenvalues by red dot and blue plus symbols, respectively. Those λ nearly on the imaginary axis correspond to the standard traveling sinusoidal waves. The amplitudes for this band of waves are hardly changed since the damping coefficients are very small. On the other hand, those λ composing a line parallel to the real axis correspond to the evanescent waves. If the absolute values of the real part of λ is large, the evanescent waves damp fast in space. In the second subplot, a transformation from λ to $\frac{\lambda}{s}$ is made. Since s is almost purely imaginary, this transformation nearly rotates the complex plane clockwise with $\frac{\pi}{2}$ angle. Obviously, the sign of some λ is varied, especially those corresponding to the evanescent waves. But a careful examination shows that if we shift the origin of the second subplot to a point fairly close to $\frac{2}{3}$, then all the eigenvalues will maintain their signs. Notice that

$$\frac{2}{3} = \frac{0.5}{1 - 0.5^2} = \frac{u_0}{c_0^2 - u_0^2} = \mu.$$

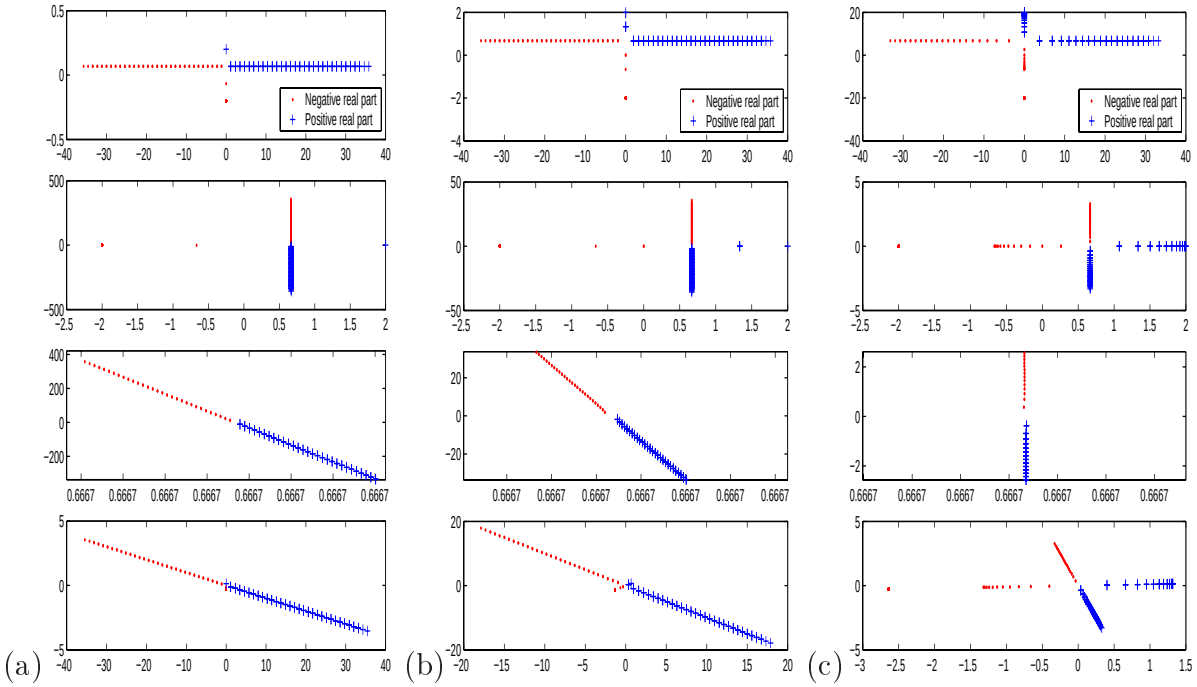


Figure 5.2: Generalized eigenvalues for the uniform base flow $u_0 = 0.5$. (a) $s = 10^{-12} + 0.1i$; (b) $s = 10^{-12} + i$; (c) $s = 10^{-12} + 10i$. First row: λ. Second row: $\frac{\lambda}{s}$. Third row: refined structure of the second. Last row: $f = \left(\frac{\lambda}{s} - \mu\right)\frac{s}{s+\alpha}$ with $\alpha = 1$.

Similar observations can be made, for different s, through Figures 5.2-a, 5.2-c. This nice property does not hold in general for the parallel flow. Figure 5.3-b demonstrates the computational results when $u_0 = M\frac{\cos^2(y/2)}{1+\sin^2(y/2)}$ with $M = 0.5$. The period L is set equal to 2π. In the second subplot, again, we see the sign of some λ is varied. But in this case, it is impossible to find a simple shift of the origin to maintain the sign, as the third subplot shows. However, by examining the second subplot which shows $\frac{\lambda}{s}$, we can find a nice property: though the red dot symbols cannot be separated from the blue plus symbols with a straight line parallel to the imaginary axis, it seems that a suitable oblique line can fulfill this purpose, see also Figure 5.3-a and 5.3-c for different s. To determine this oblique line, we can make two successive transformations: first shift the origin to some point μ, then rotate the complex plane $\frac{\lambda}{s} - \mu$ by an appropriate angle. Of course, we do not expect μ to rely on the choice of s. Besides, the rotating transformation should be easy to handle. Stimulated by the work of Hagstrom [**42**] and Hu [**54**], for the parallel flow with constant density,

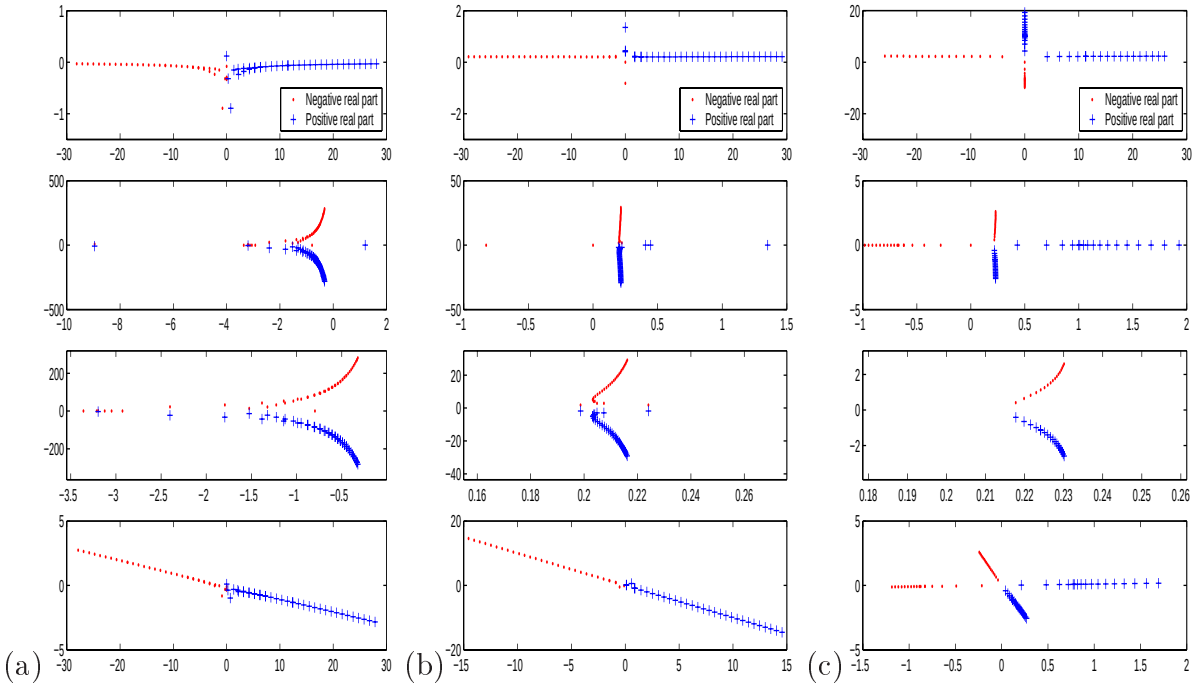


Figure 5.3: Generalized eigenvalues for nonuniform base flow $u_0 = 0.5\frac{\cos^2(y/2)}{1+\sin^2(y/2)}$. (a) $s = 10^{-12} + 0.1i$; (b) $s = 10^{-12} + i$; (c) $s = 10^{-12} + 10i$. First row: λ. Second row: $\frac{\lambda}{s}$. Third row: refined structure of the second. Last row: $f = \left(\frac{\lambda}{s} - \mu\right)\frac{s}{s+\alpha}$ with $\alpha = 1$.

we make transformation analogous to (5.13) as

$$\lambda \longrightarrow f = \left(\frac{\lambda}{s} - \mu\right)\frac{s}{s+\alpha}, \tag{5.21}$$

with

$$\mu = \frac{u_m}{c_0^2 - u_m^2},\ u_m = \frac{1}{L}\int_0^L u(y)dy,\ \alpha = \frac{2\pi c_0}{L}. \tag{5.22}$$

The fourth row of subplots in Figure 5.3 shows the results after transformation. All the eigenvalues λ indeed maintain their sign.

We can also detect the effect of this phase shift parameter α for the uniform flow. The fourth row in Figure 5.2 shows the results. We see that α typically increases the damping rates of the evanescent waves, though it sacrifices those of the traveling waves a little. This parameter was first introduced in [**63**] for electromagnetic problems.

We should confess that though the choice of μ and α in (5.22) is valid in many cases, in some exceptional situations when the amplitude of s is relatively small, we do find from our numerical computations that several few generalized eigenvalues

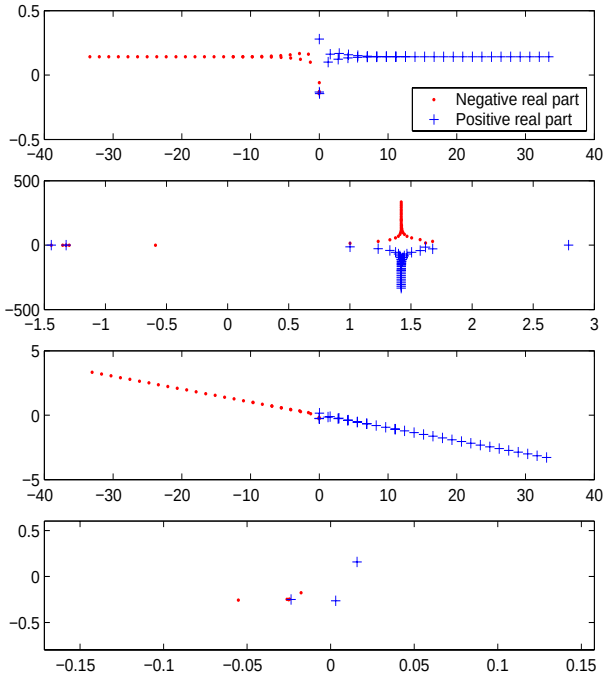


Figure 5.4: Generalized eigenvalues for parallel flow $u_0 = 0.4\frac{\cos^2(y/2)}{1+\sin^2(y/2)} + 0.5$. $s = 10^{-12} + 0.1i$. Top: λ. Second: $\frac{\lambda}{s}$. Third: $f = \left(\frac{\lambda}{s} - \mu\right)\frac{s}{s+\alpha}$ with $\alpha = 1$. Bottom: refined structure of the third subplot.

cannot maintain their sign after the transformation (5.21). See Figure 5.4 for an example. But in these situations, the real part of f is small, which implies a slow increase of amplitude in the PML layers. Since the thickness of the PML layers is usually made small, and the characteristic boundary condition is used at the PML boundaries, it is reasonable to expect a small error from these exceptional wave modes.

In the above analysis, we assume that the parallel flow is periodic in y-direction with period L. In the real applications, either for the ducted flow or for the open flow, this assumption is unreasonable. In these cases, we can still make the transformation (5.21) with μ and α determined in (5.22). But the period L should be replaced with the characteristic wave length, which is usually a problem-dependent parameter, and has to be determined by trial and error.

After μ and α are determined, the argument of deriving the PML equations for the uniform flow can be made analogously for the parallel flow. We omit this discussion.

The final PML equations formulated with the conservative variables read

$$\tilde{q}_t + A\tilde{q}_x + B\tilde{q}_y + B^*\tilde{q} + \sigma\mu A\tilde{q} + \sigma w = 0, \tag{5.23}$$

$$w_t + A\tilde{q}_x + (\alpha+\sigma)w + (\alpha+\sigma)\mu A\tilde{q} = 0. \tag{5.24}$$

5.3 PML formulations for the nonlinear Euler equations

In the last section, we discussed the PML for the linearized Euler equations. The key point is to determine the parameters μ and α. Remember what we are trying to solve is the nonlinear Euler equations. If we insist on using the linearized Euler equations in the PML absorbing layers, we have to solve a mixed-type problem: the nonlinear Euler equations in the physical domain and the linearized Euler equations in the PML absorbing layers, combined with the artificial boundaries. In this case, the interface condition at the artificial boundaries has to be considered carefully. From the computational point of view, this is troublesome. In the following, we present a simple idea to solve this problem.

First we consider the PML equations (5.14)-(5.15) for the uniform flow. Replacing $\tilde{q}$ with $q - q_0$ we have

$$q_t + Aq_x + Bq_y + \sigma\mu A(q - q_0) + \sigma w = 0, \tag{5.25}$$

$$w_t + Aq_x + v_0 w_y + (\alpha+\sigma)\mu A(q - q_0) + (\alpha+\sigma)w = 0. \tag{5.26}$$

Since

$$f(q) \approx f(q_0) + A(q - q_0),\ g(q) \approx g(q_0) + B(q - q_0)$$

within the first order accuracy, by replacing Aq_x with $f(q)_x$, Bq_y with $g(q)_y$ and $A(q - q_0)$ with $f(q) - f(q_0)$, we derive a first order approximation of the equations (5.25)-(5.26)

$$q_t + f(q)_x + g(q)_y + \sigma\mu(f(q) - f(q_0)) + \sigma w = 0, \tag{5.27}$$

$$w_t + f(q)_x + v_0 w_y + (\alpha+\sigma)\mu(f(q) - f(q_0)) + (\alpha+\sigma)w = 0. \tag{5.28}$$

If we extend the definition of σ to the physical domain with zero, and consider a profile of σ sufficiently smooth, then the equations (5.27)-(5.28) are naturally compatible with the original nonlinear Euler equations. Thus the equations (5.27)-(5.28) are valid in the whole computational domain, including both the physical domain

and the PML absorbing layers. Besides, as the PML equations (5.14)-(5.15) are derived from the linearized Euler equations, which itself is a first order approximation of the nonlinear Euler equations, the solution of the nonlinear PML equations (5.27)-(5.28) being restricted to the physical domain is a first order approximation of the original nonlinear Euler equations (5.1). The PML equations (5.23)-(5.24) for the parallel flow can be modified analogously. The resulting nonlinear PML equations read

$$q_t + f(q)_x + g(q)_y + \sigma\mu(f(q) - f(q_0(\infty, y, t))) + \sigma w = 0, \tag{5.29}$$
$$w_t + f(q)_x + (\alpha + \sigma)\mu(f(q) - f(q_0(\infty, y, t))) + (\alpha + \sigma)w = 0. \tag{5.30}$$

These equations can be taken as a special case of (5.27)-(5.28) for $v_0 = 0$. Thus in both cases, the nonlinear PML equations can be formulated in a unified form (5.27)-(5.28). The only point to keep in mind is that the mean state variable q_0 should be valued at infinity in the correct direction.

For open flow problems, we enclose the physical domain with the PML absorbing layers in both two directions. A straightforward extension of the nonlinear PML equations (5.27)-(5.28) to two dimensions is

$$q_t + f(q)_x + g(q)_y + \sigma_x\mu_x(f(q) - f(q_\infty)) + \sigma_y\mu_y(g(q) - g(q_\infty)) + \sigma_x w_1 + \sigma_y w_2 = 0, \tag{5.31}$$
$$w_{1,t} + f(q)_x + v_\infty w_{1,y} + (\alpha_x + \sigma_x)w_1 + (\alpha_x + \sigma_x)\mu_x(f(q) - f(q_\infty)) = 0, \tag{5.32}$$
$$w_{2,t} + u_\infty w_{2,x} + g(q)_y + (\alpha_y + \sigma_y)w_2 + (\alpha_y + \sigma_y)\mu_y(g(q) - g(q_\infty)) = 0. \tag{5.33}$$

Here, μ_x, α_x and μ_y, α_y are determined by (5.22) in the x- and y-directions, respectively. Unfortunately, this seemingly natural treatment suffers from some unstable fast-growing wave modes, even for the uniform flow. This instability has been noticed in [**6**]. In Figure 5.5, we show the maximal growth rate obtained by the numerical computation of all wave modes of (5.31)-(5.33), with the fixed wave numbers k_x and k_y. We can see that as the wave numbers increase, the maximal growth rate grows rapidly. Since the positive growth rate only appears when both σ_x and σ_y are positive, we conclude that this instability only appears at the corner region of the PML domain. This is verified by our numerical computations, see Figure 5.10.

To suppress these instabilities, we simply add an artificial damping term

$$(\alpha_x + \alpha_y)\frac{\sigma_x\sigma_y}{\sigma_x^m + \sigma_y^m}(q - q_\infty)$$

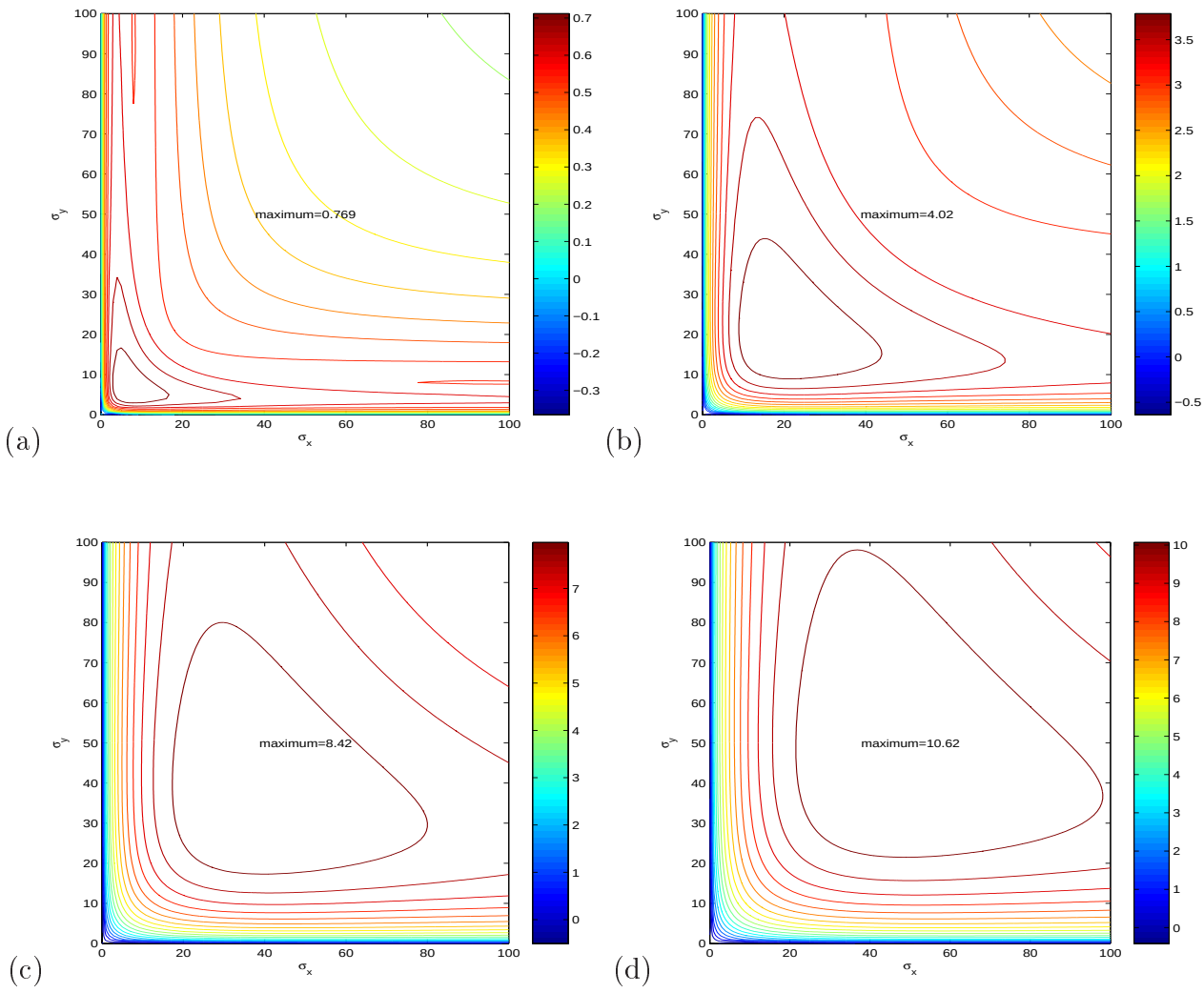


Figure 5.5: Maximal growth rates. (a) $k_x = k_y = 10$; (b) $k_x = k_y = 40$; (c) $k_x = k_y = 80$; (d) $k_x = k_y = 100$.

to equation (5.31). Here σ_x^m and σ_y^m are the maximal values of σ_x and σ_y, respectively. The final nonlinear PML equations for the open flow problems are then

$$q_t + f(q)_x + g(q)_y + \sigma_x\mu_x(f(q) - f(q_\infty)) + \sigma_y\mu_y(g(q) - g(q_\infty)) + \sigma_x w_1 + \sigma_y w_2 + (\alpha_x + \alpha_y)\frac{\sigma_x\sigma_y}{\sigma_x^m + \sigma_y^m}(q - q_\infty) = 0, \quad (5.34)$$

$$w_{1,t} + f(q)_x + v_\infty w_{1,y} + (\alpha_x + \sigma_x)w_1 + (\alpha_x + \sigma_x)\mu_x(f(q) - f(q_\infty)) = 0, \quad (5.35)$$

$$w_{2,t} + u_\infty w_{2,x} + g(q)_y + (\alpha_y + \sigma_y)w_2 + (\alpha_y + \sigma_y)\mu_y(g(q) - g(q_\infty)) = 0. \quad (5.36)$$

Our numerical tests show that this treatment works well.

5.4 Solution strategies

In this section we discuss solution strategies of solving the nonlinear PML equations (5.34)-(5.36). This is a hyperbolic system with a local source term. When either σ_x or σ_y turns large, the source term becomes stiff. In this case, any explicit numerical scheme would suffer from the strict constraint on the time step size.

A naive implicit discretization of the time variable is absolutely not a good resolution. This is because any kind of spatial discretization would result in a scheme in which some nonlinear algebraic systems with a large number of unknowns need to be solved during the time advancing. To overcome this difficulty, we can resort to the Implicit-Explicit Runge-Kutta (IERK) semi-discretization method.

The IERK method is composed of two steps. First we should perform the spatial discretization for the equations (5.34)-(5.36). There are many choices which can serve this purpose. Since the nonlinear Euler equations in general admit discontinuous shock wave solution, we prefer the WENO-type schemes. Though high-order versions can be considered without any technical difficulty, in this paper we only employ a second order finite volume scheme based on the linear reconstruction with the minmod limiter. On the cell interface, the flux is computed with Roe's approximate Riemann solver. More details can be found in the book of Leveque [**64**]. For the source term, we simply take its cell average as the function value of the cell average state, which holds within second order accuracy.

The second step of the IERK method requires solving a large ODE system resulting from the first step, say,

$$U_t = H(U) + S(U), \tag{5.37}$$

where H comes from the hyperbolic part and S from the stiff source term. U denotes the vector composed of all cell averages of q, w_1 and w_2. Suppose Δt is the time step and the numerical solution at the n-th step is already obtained as U_n. The IMEX Runge-Kutta approximation at the $n+1$-th step is realized by

$$U^i = U_n + \Delta t \sum_{j=1}^{i-1} \tilde{a}_{ij} H(U^j) + \Delta t \sum_{j=1}^{\nu} a_{ij} S(U^j),\ i = 1, \cdots, \nu, \tag{5.38}$$

$$U_{n+1} = U_n + \Delta t \sum_{i=1}^{\nu} \tilde{w}_i H(U^i) + \Delta t \sum_{i=1}^{\nu} w_i S(U^i). \tag{5.39}$$

Here, the matrices $\tilde{A} = (\tilde{a}_{ij})$, $\tilde{a}_{ij} = 0$ for $j \geq i$ and $A = (a_{ij})$ are $\nu \times \nu$ matrices. $\tilde{A}$ and $\tilde{W} = (\tilde{w}_i)$ correspond to an explicit Runge-Kutta scheme, while A

and $W = (w_i)$ correspond to an implicit Runge-Kutta scheme. To efficiently solve the nonlinear algebraic system corresponding to the implicit part, it is natural to consider diagonally implicit Runge-Kutta (DIRK) schemes for the source term, i.e., $a_{ij} = 0$ for $j > i$.

For a second order approximation, we use the following two-stage scheme

$$U^1 = U_n + \chi \Delta t S(U^1), \tag{5.40}$$

$$U^2 = U_n + \Delta t H(U^1) + (1 - 2\chi)\Delta t S(U^1) + \chi \Delta t S(U^2), \tag{5.41}$$

$$U_{n+1} = U_n + \frac{1}{2}\Delta t \left(H(U^1) + H(U^2) + S(U^1) + S(U^2) \right), \tag{5.42}$$

with $\chi = 1 - \frac{\sqrt{2}}{2}$. This is also an L-stable TVD scheme. The readers are referred to **[9, 68]** for more detail.

Notice that we still need to solve an algebraic system with a large number of unknowns in (5.40) or (5.41). But unlike the naive implicit scheme such as the backward Euler, this large algebraic system can be decoupled to a lot of small algebraic systems with only several few unknowns related to each specific cell. In fact, in our scheme at most 12 cell average variables are involved in each small algebraic system, and the number can be further reduced to 4 by some basic calculus. For example, let us consider the implicit step (5.40). On a cell denoted by symbol *, we have to solve

$$\begin{aligned} &\frac{q^1_* - q_*}{\chi \Delta t} + \sigma_{x,*}\mu_x(f(q^1_*) - f(q_{\infty,*})) + \sigma_{y,*}\mu_y(g(q^1_*) - g(q_{\infty,*})) \\ &\qquad + \sigma_{x,*} w^1_{1,*} + \sigma_{y,*} w^1_{2,*} + (\alpha_x + \alpha_y)\frac{\sigma_{x,*}\sigma_{y,*}}{\sigma^m_x + \sigma^m_y}(q^1_* - q_\infty) = 0, \end{aligned} \tag{5.43}$$

$$\frac{w^1_{1,*} - w_{1,*}}{\chi \Delta t} + (\alpha_x + \sigma_{x,*}) w^1_{1,*} + (\alpha_x + \sigma_{x,*})\mu_x(f(q^1_*) - f(q_{\infty,*})) = 0, \tag{5.44}$$

$$\frac{w^1_{2,*} - w_{2,*}}{\chi \Delta t} + (\alpha_y + \sigma_{y,*}) w^1_{2,*} + (\alpha_y + \sigma_{y,*})\mu_y(g(q^1_*) - g(q_{\infty,*})) = 0. \tag{5.45}$$

Here, for conciseness of notations, we omit the subfix n for the field variables q, w_1 and w_2. (5.43)-(5.45) is a nonlinear algebraic system with 12 unknowns. From (5.44) and (5.45), $w^1_{1,*}$ and $w^1_{2,*}$ can be expressed with q^1_* as

$$w^1_{1,*} = \frac{1}{1 + \chi \Delta t(\alpha_x + \sigma_{x,*})} \left(w_{1,*} - \chi \Delta t(\alpha_x + \sigma_{x,*})\mu_x(f(q^1_*) - f(q_{\infty,*})) \right), \tag{5.46}$$

$$w^1_{2,*} = \frac{1}{1 + \chi \Delta t(\alpha_y + \sigma_{y,*})} \left(w_{2,*} - \chi \Delta t(\alpha_y + \sigma_{y,*})\mu_y(g(q^1_*) - g(q_{\infty,*})) \right). \tag{5.47}$$

Substituting (5.46) and (5.47) into (5.43) we get

$$\frac{q_*^1}{\chi\Delta t}+\frac{\sigma_{x,*}\mu_x}{1+\chi\Delta t(\alpha_x+\sigma_{x,*})}(f(q_*^1)-f(q_{\infty,*}))+\frac{\sigma_{y,*}\mu_y}{1+\chi\Delta t(\alpha_y+\sigma_{y,*})}(g(q_*^1)-g(q_{\infty,*}))$$
$$=\frac{q_*}{\chi\Delta t}-\frac{\sigma_{x,*}w_{1,*}}{1+\chi\Delta t(\alpha_x+\sigma_{x,*})}-\frac{\sigma_{y,*}w_{2,*}}{1+\chi\Delta t(\alpha_y+\sigma_{y,*})}-(\alpha_x+\alpha_y)\frac{\sigma_{x,*}\sigma_{y,*}}{\sigma_x^m+\sigma_y^m}(q_*^1-q_\infty). \tag{5.48}$$

Only 4 unknowns are involved in (5.48), which can be solved with Newton iterations. To ensure a nonsingular Jacobian in each iteration, we have to make constraints on the absorbing functions σ_x and σ_y. Suppose c_{max} is the maximal sound speed. Then a sufficient condition for nonsingular Jacobians could be

$$\frac{1}{\chi\Delta t}\geq 2c_{max}(\sigma_x^m\mu_x+\sigma_y^m\mu_y).$$

Since c_{max} is unknown, in the computation we replace c_{max} with the maximal c_∞ and set

$$\sigma_x^m=\sigma_y^m=\sigma^m=\frac{1}{2(\mu_x+\mu_y)\chi\Delta t\max c_\infty}. \tag{5.49}$$

We still need to determine the profile of the absorbing functions and the thickness D of the PML domain at each artificial boundary. In principle, the bigger is D, the better is the approximation. But considering the computational cost, under the precondition of accuracy, D should be reduced as small as possible. From (5.12) and (5.13), we see that on the interval $[x_0, x_0+D]$, the damping rates for all modal solutions depend, in addition to f, on a common factor $\int_{x_0}^{x_0+D}\sigma(z)dz$. If we set

$$\sigma=\sigma^m\left(\frac{x-x_0}{D}\right)^N,\ x\in[x_0,x_0+D],$$

then

$$\int_{x_0}^{x_0+D}\sigma(z)dz=\frac{D\sigma^m}{N+1}.$$

Denoting the above quantity by N_0 and consulting (5.49), we have

$$D=\left[2(N+1)N_0(\mu_x+\mu_y)\chi\max c_\infty\right]\Delta t.$$

This implies that for a specific problem, the number of grid points in the PML domain is always fixed with the proposed method. In the real implementation, for the accuracy requirement, we set

$$D=\max\left(\left[2(N+1)N_0(\mu_x+\mu_y)\chi\max c_\infty\right],10\right)\Delta t, \tag{5.50}$$

which means at least 10 grid points are used in the PML absorbing layers.
For easy reference, we list the main steps of solving the nonlinear open flow problems with the equations (5.34)-(5.36):

1. fix the physical part of the computational domain by introducing artificial boundaries;
2. prescribe reasonable mean flow at each artificial boundary;
3. determine the characteristic wave length, thus α and μ by equation (5.22);
4. determine the spatial step sizes Δx and Δy, thus Δt under the CFL constraint;
5. determine σ^m with equation (5.49);
6. determine the profile parameter N and N_0 (3 and 15 used in our computations);
7. determine the width of the PML domain D by (5.50) at each artificial boundary;
8. solve the nonlinear PML equations (5.34)-(5.36) with the scheme (5.40)-(5.42).

Other flow problems can be considered analogously.

5.5 Numerical tests

In this section, we present two numerical tests, dealing with an open uniform flow and a ducted parallel flow respectively.

5.5.1 Open uniform flow

Suppose that the base state is (ρ_0, u_0, v_0, p_0) with $\rho_0 = 1$ and $p_0 = \frac{1}{\gamma}$. At the initial time the pressure field is set to

$$p = \frac{1}{\gamma} + Ae^{-\ln(2)(x^2+y^2)/0.2^2}$$

with $A = 1$. The special case with $u_0 = 0.5$ and $v_0 = 0$ has been considered in the work of Hu [**55**], in which the physical domain is set as $[-1, 1] \times [-1, 1]$ with $\Delta x = \Delta y = 0.02$. By (5.22), $\mu_x = \frac{2}{3}$, $\mu_y = 0$, and $\alpha_x = \alpha_y = 2\pi$. The time step is $\Delta t = 0.005$ satisfying the CFL constraint. According to (5.49), $\sigma^m \approx 330.6$, and

from (5.50), twelve grid points are needed in the PML domain. Thus the computational domain is $[-1.24, 1.24] \times [-1.24, 1.24]$.

To evaluate the quality of numerical solutions, we take the reference solutions as those computed on a large domain $[-4, 4] \times [-4, 4]$ with the same mesh size. The errors for different field variables are depicted in Figure 5.6. For comparison, we also plot the errors of the numerical solutions with the characteristic boundary conditions. We see that the relative L_1 errors (except for the v component since $v_0 = 0$) are less than 0.5 percent in the time interval $[0, 5]$. The performance of the PML is

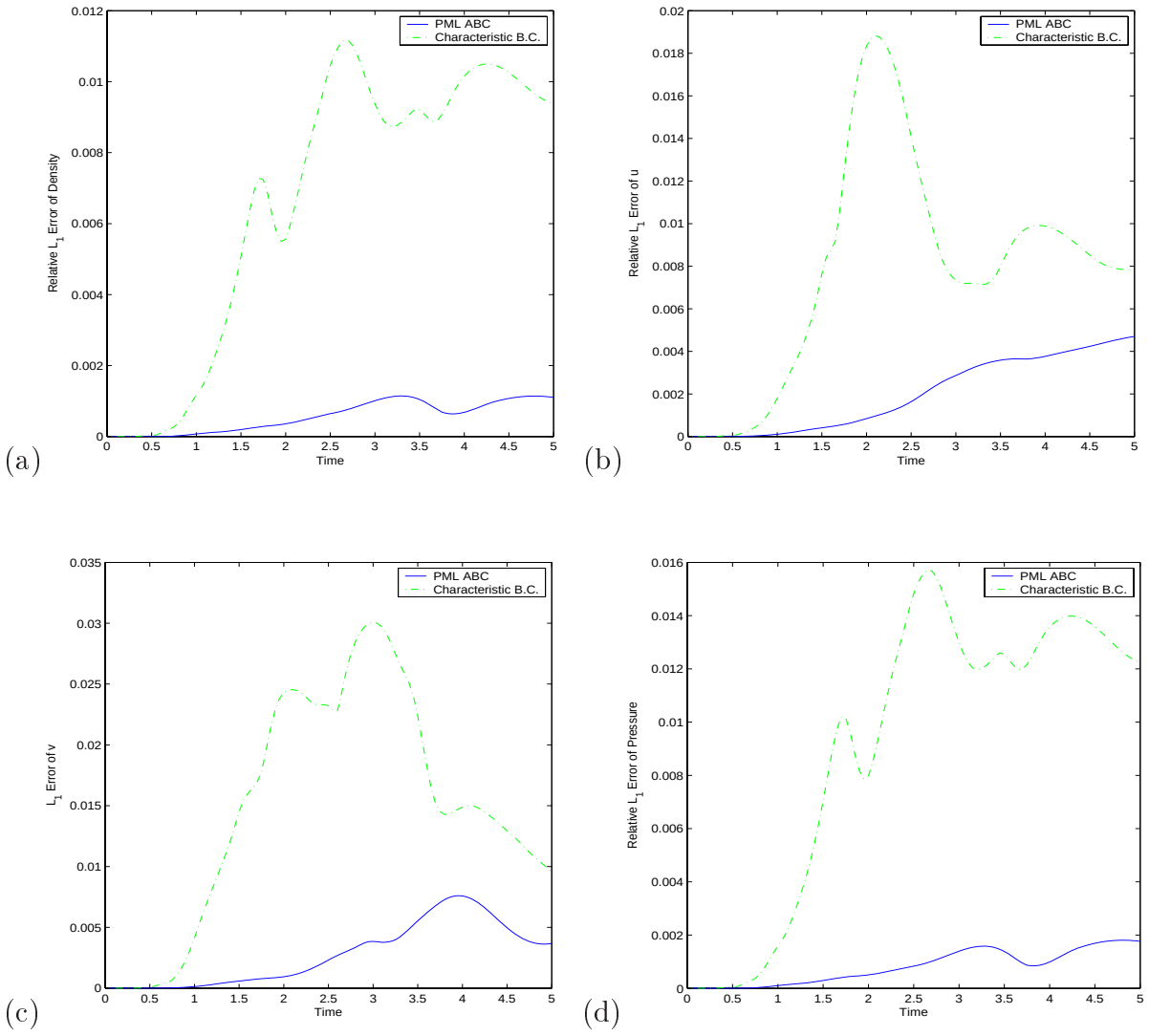


Figure 5.6: Comparison between the PML ABC and the characteristic boundary condition. Uniform flow. $u_0 = 0.5$, $v_0 = 0$.

much better than that of the characteristic boundary conditions. Figure 5.7 shows the computed pressure field of the uniform flow $(u_0, v_0) = (0.5, 0)$ at $t = 0.5, 1.0$ and 1.5. Little reflections are observed.

We also consider the uniform flow but with an oblique direction, i.e., $(u_0, v_0) = (\frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4})$. In this case, $\mu_x = \mu_y = \frac{\sqrt{2}}{4-\sqrt{2}}$ and $\sigma^m \approx 272.7$. Twelve grid points are used in the PML domain. The numerical results are shown in Figure 5.8 and Figure 5.9.

In Section 3 we analyzed that if no additional treatment is performed in the corner domain, the nonlinear PML equations are unstable. In Figure 5.10, we illustrate the numerical solution at $t = 2.8$ without corner modification. Peaks form in the corners. After this time point, the computation breaks down immediately since the assumption of subsonic flow is violated. We should remark that this instability cannot be removed by reducing the time step, which implies that it does not come from our proposed numerical scheme.

In fact, after a dissipative term is added into (5.43), the observed instability disappears, and long time computations can be carried out. Figure 5.11 shows the results at $t = 100$. It strongly suggests that the numerical solutions restricted to the physical domain $[-1, 1] \times [-1, 1]$ converge to the exact solutions as time goes to infinity.

5.5.2 Ducted parallel flow

We consider a y-periodic parallel flow with

$$\rho_0 = 1,\ u_0 = 0.5\frac{\sin^2 \pi y}{1 + \cos^2 \pi y},\ v_0 = 0,\ p_0 = \frac{1}{\gamma}.$$

Since the period is 1, we can limit this problem in $\mathcal{R} \times [0, 1]$. At the initial time, a disturbance is added to the pressure field

$$\tilde{p} = p - p_0 = e^{-\ln(2)(x^2+(y-0.5)^2)/0.1^2}.$$

This disturbance is well-supported in $[-1, 1] \times [0, 1]$. Thus we take it as the physical domain. From (5.22), we have $\mu = \frac{1}{2}$ and $\alpha = 2\pi$. We set $\Delta x = \Delta y = 0.01$ and $\Delta t = 0.0025$. By (5.49), we have $\sigma^m = 3155.6$, and by (5.50), we know ten grid points are needed in the x-direction of the PML domain. The computational domain is thus $[-1.1, 1.1] \times [0, 1]$.

The numerical errors are depicted in Figure 5.12. The relative errors are typically below 0.5 percent. The advantage of our PML ABC over the characteristic boundary condition is obvious.

To examine the stability property, we compute the solution at time $t = 10$ and $t = 100$. The results are shown in Figure 5.13 and Figure 5.14, respectively. It suggests that as time goes to infinity, the solution converges to a state which is close to the steady state solution.

5.6 Conclusion

Based on the work of Hagstrom [**42**] and Hu [**55**], we have presented a new PML formulation for the nonlinear Euler equations. Both uniform flow and nonuniform but parallel flow have been considered. One of the main advantage of our formulation lies in the autonomic form of its hyperbolic part. The technique for the nonlinear conservation laws can be used directly. For the flow problems with shock waves, this treatment is crucial.

One of the key points to successfully use the PML is to find out a way of damping all wave modes in their trajectories in the PML layers. This has been solved completely by Hagstrom for the uniform flow, but for the parallel flow, from theoretical perspective, it is still an unsolved problem. As revealed in Hagstrom and Nazarov [**43**], only μ cannot ensure the damping of all wave modes. In this paper we have made some progress in proposing an approach of determining μ and α to damp all the wave modes for the parallel flow. Numerical tests have shown that for most wave modes the proposed values can damp them in the PML layers. Exceptional cases appear when the wave frequency is relatively small. But in these cases, those wave modes have a small amplifying factor, and the resulting error is expected to be reasonably small.

We have also analyzed the corner instability of the PML equations. This is much analogous to the local boundary conditions for which the corner compatibility conditions play an important role in the stability of the truncated problem. Our numerical tests have shown that with a suitable dissipative term added in the corner domain, this instability can be removed.

The PML equations are usually hyperbolic systems with source term. This source term becomes stiff when σ^m, the maximal value of the absorption coefficient σ, is large. Thus, to use an explicit numerical scheme, we have to use a fairly small time step. Alternatively, we can reduce σ^m to use a moderate time step. But in this case, the PML absorbing layer has to be broadened, which results in an increased computational cost. To resolve this contradict, we have used the Implicit-Explicit

Runge-Kutta semi-discretization method. Since the source term in the PML formulation is purely local, the implicit part with this method can be solved efficiently. Finally, we should remark that the discussions in this paper for the nonlinear Euler equations can be also applied to some other nonlinear conservation laws. For example, the PML formulation for the shallow water equations can be obtained in a straightforward manner.

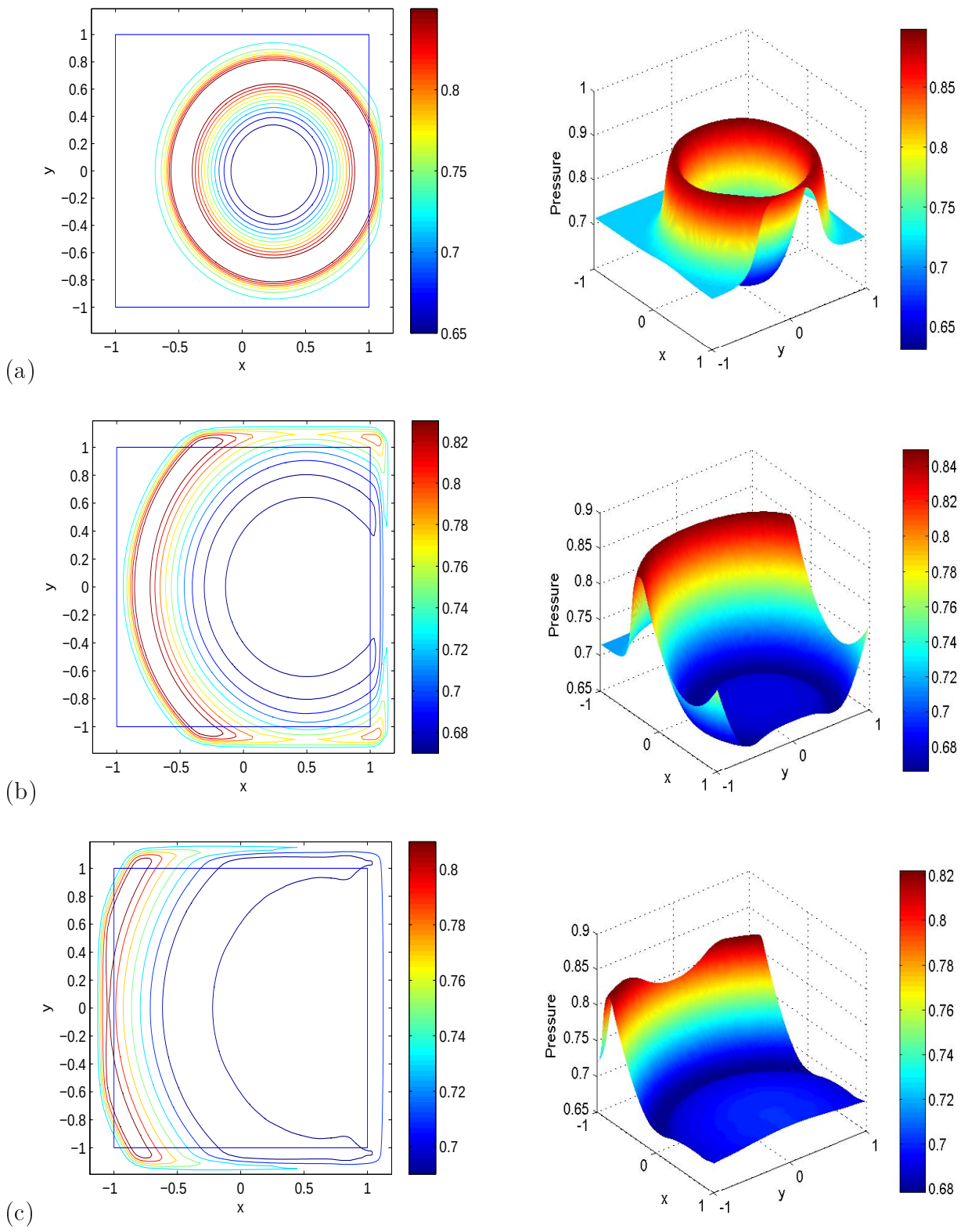


Figure 5.7: Computed pressure with PML ABC at different time points for the uniform flow $(\rho_0, u_0, v_0, p_0) = (1, 0.5, 0, 1/\gamma)$. (a) $t = 0.5$; (b) $t = 1.0$; (c) $t = 1.5$.

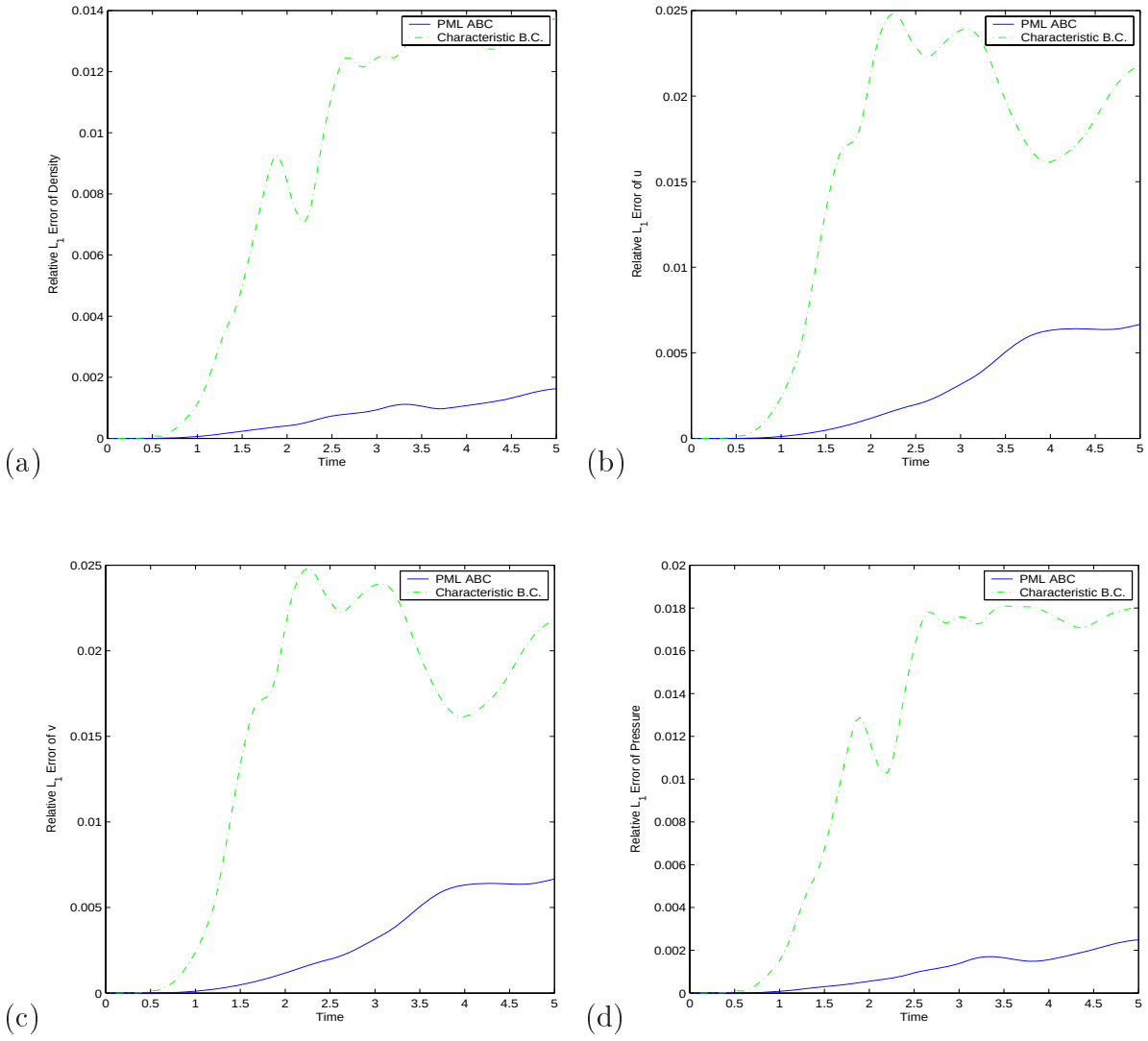


Figure 5.8: Comparison between the PML ABC and the characteristic boundary condition. Uniform oblique flow. $u_0 = v_0 = \frac{\sqrt{2}}{4}$.

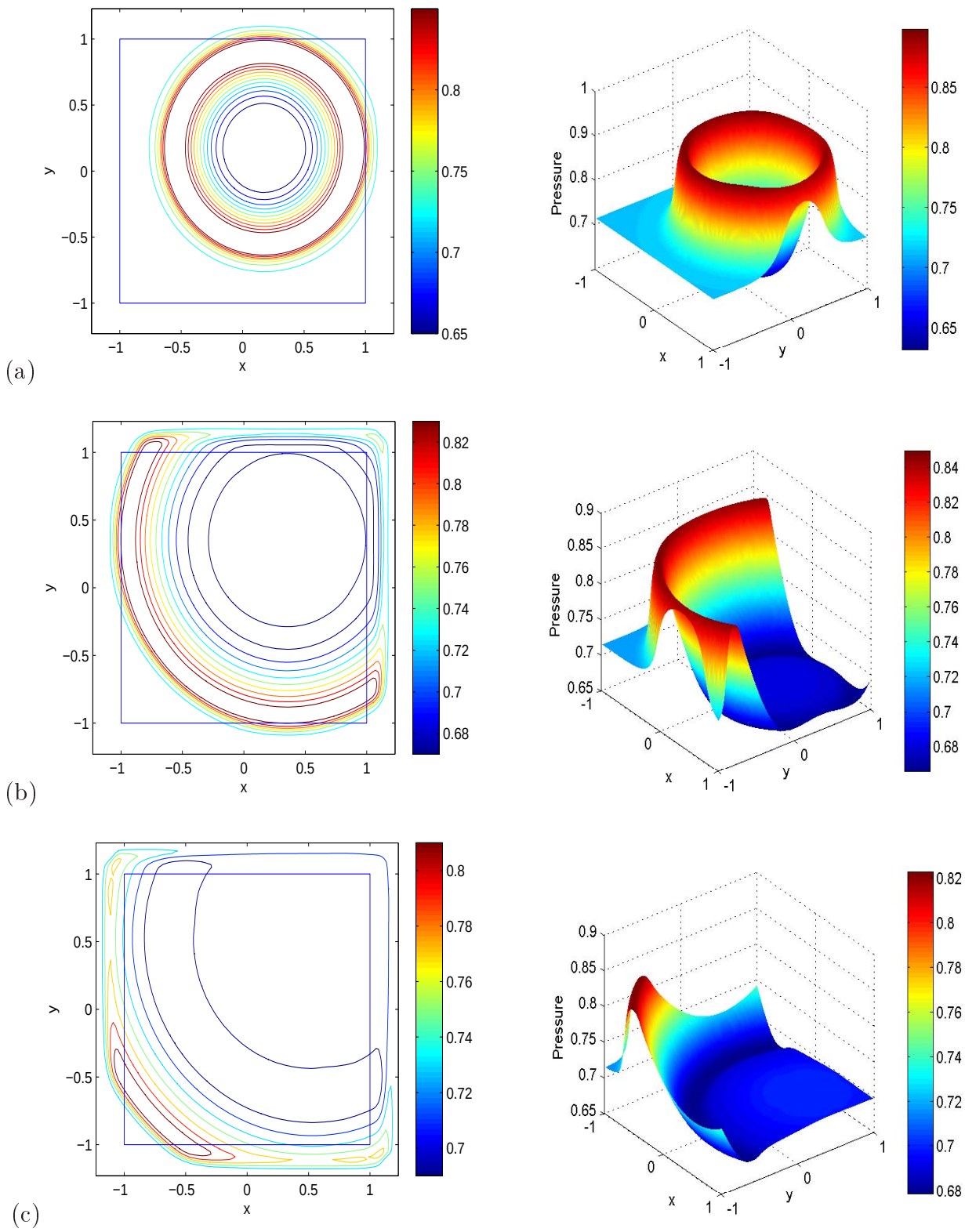


Figure 5.9: Computed pressure with PML ABC at different time points for the uniform flow $(\rho_0, u_0, v_0, p_0) = (1, \frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}, 1/\gamma)$. (a) $t = 0.5$; (b) $t = 1.0$; (c) $t = 1.5$.

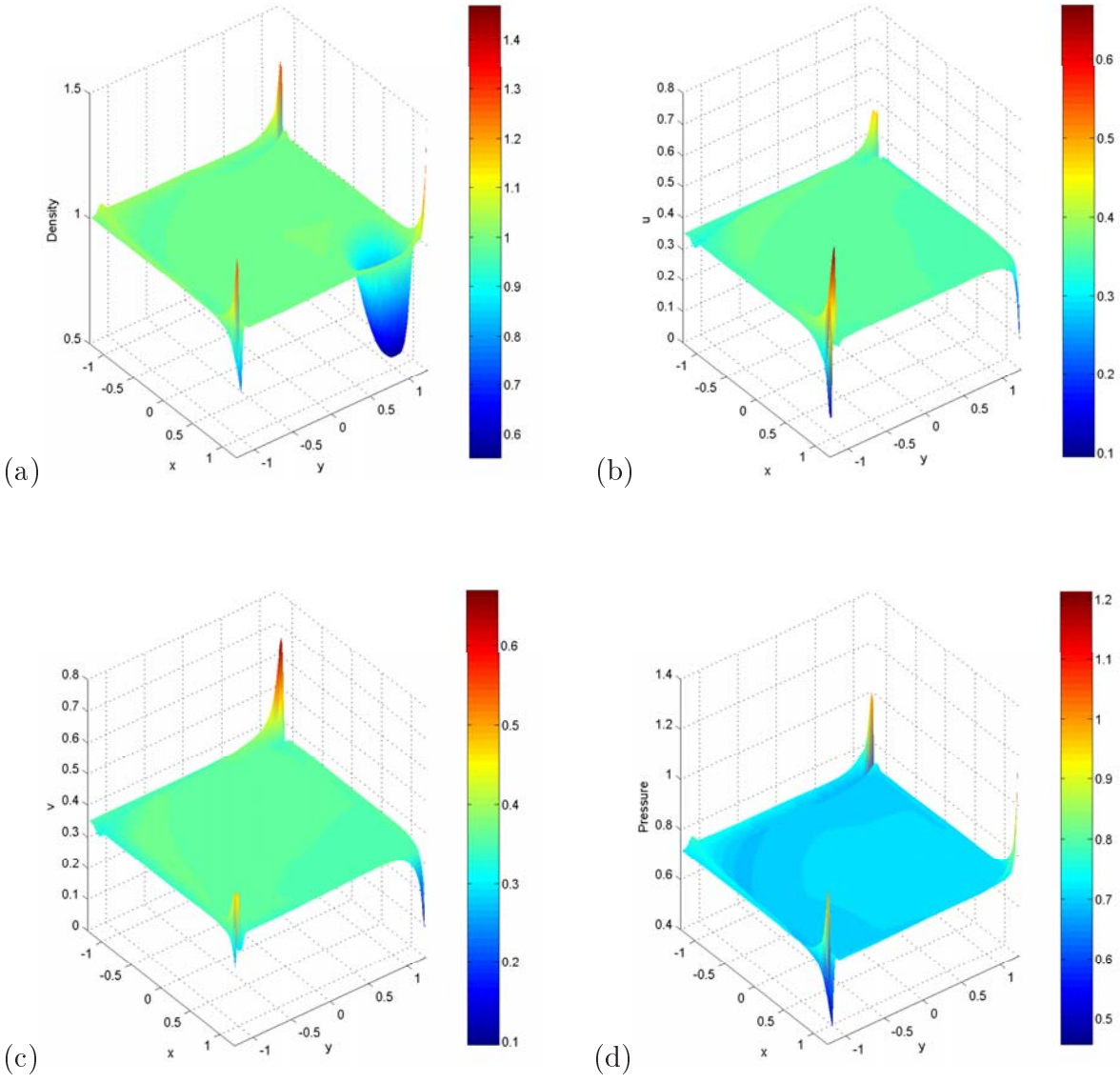
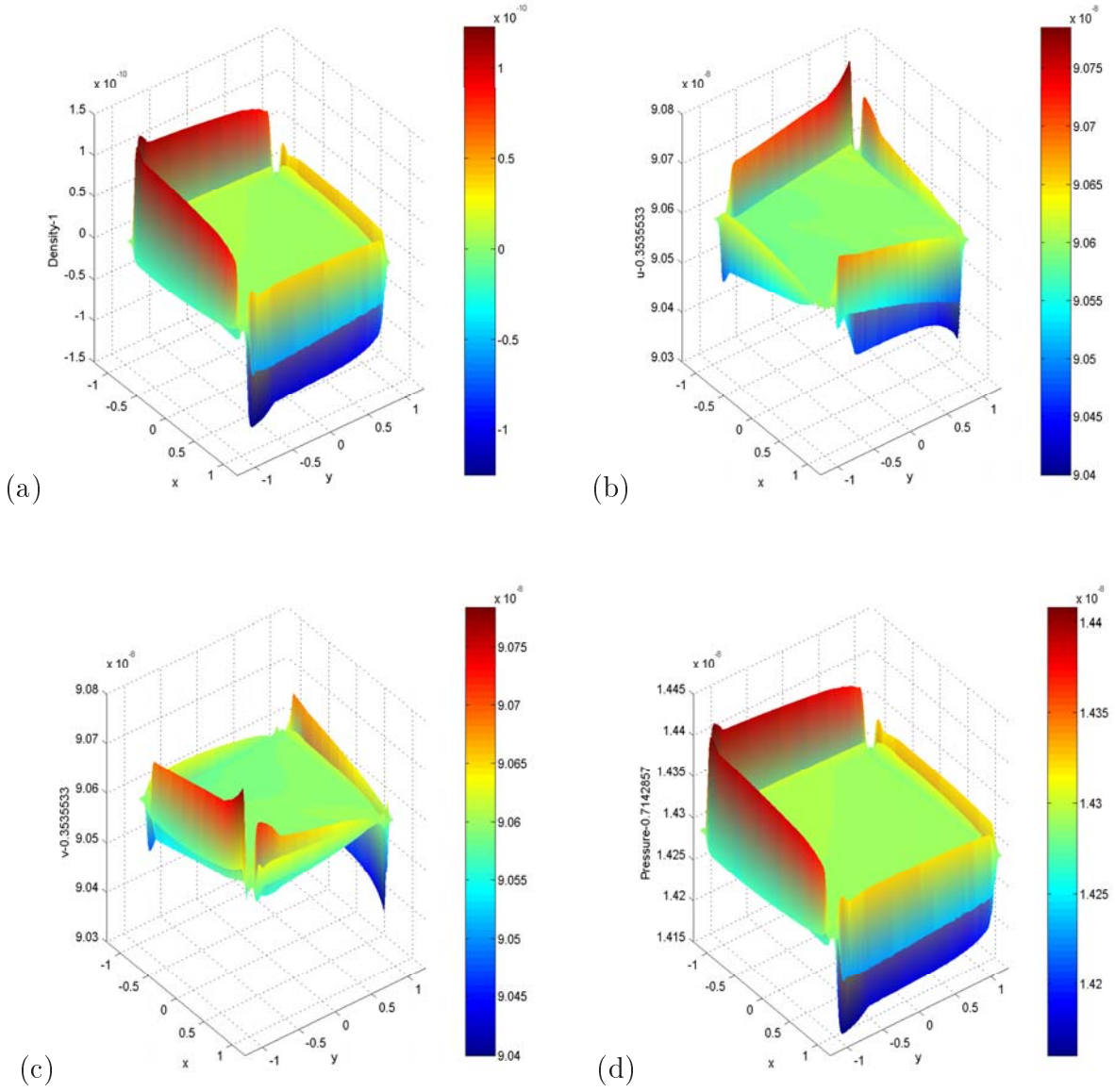


Figure 5.10: Numerical solution at $t = 2.8$ without corner modification. Oblique flow $u_0 = v_0 = \frac{\sqrt{2}}{4}$. (a) Density; (b) u; (c) v; (d) Pressure.



(a) (b) (c) (d)

Figure 5.11: Difference between the numerical solution and the steady state solution at $t = 100$. Oblique flow. $u_0 = v_0 = \frac{\sqrt{2}}{4}$. (a) Density; (b) u; (c) v; (d) Pressure.

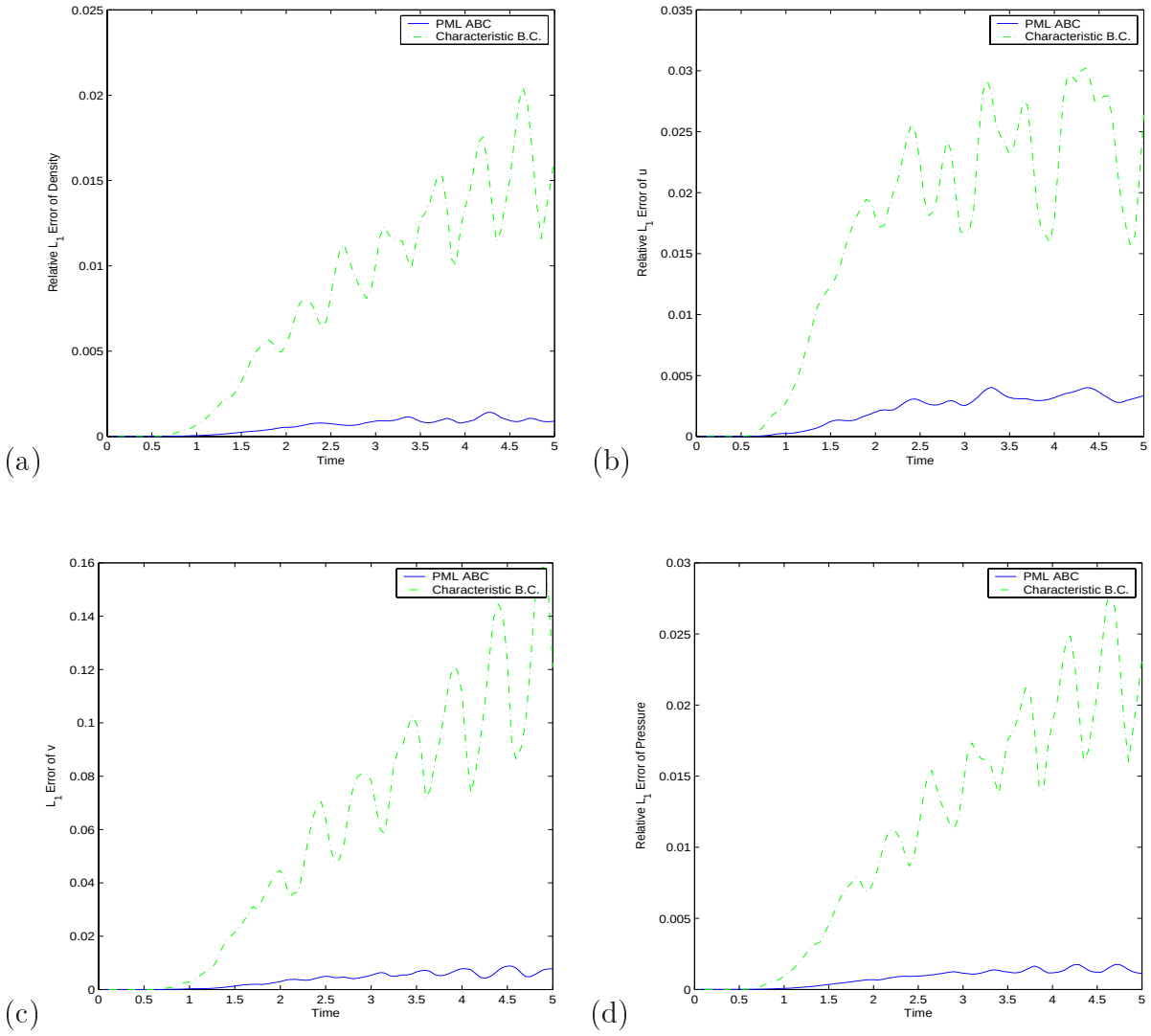


Figure 5.12: Comparison between the PML ABC and the characteristic boundary condition. (a) Density; (b) u; (c) v; (d) Pressure.

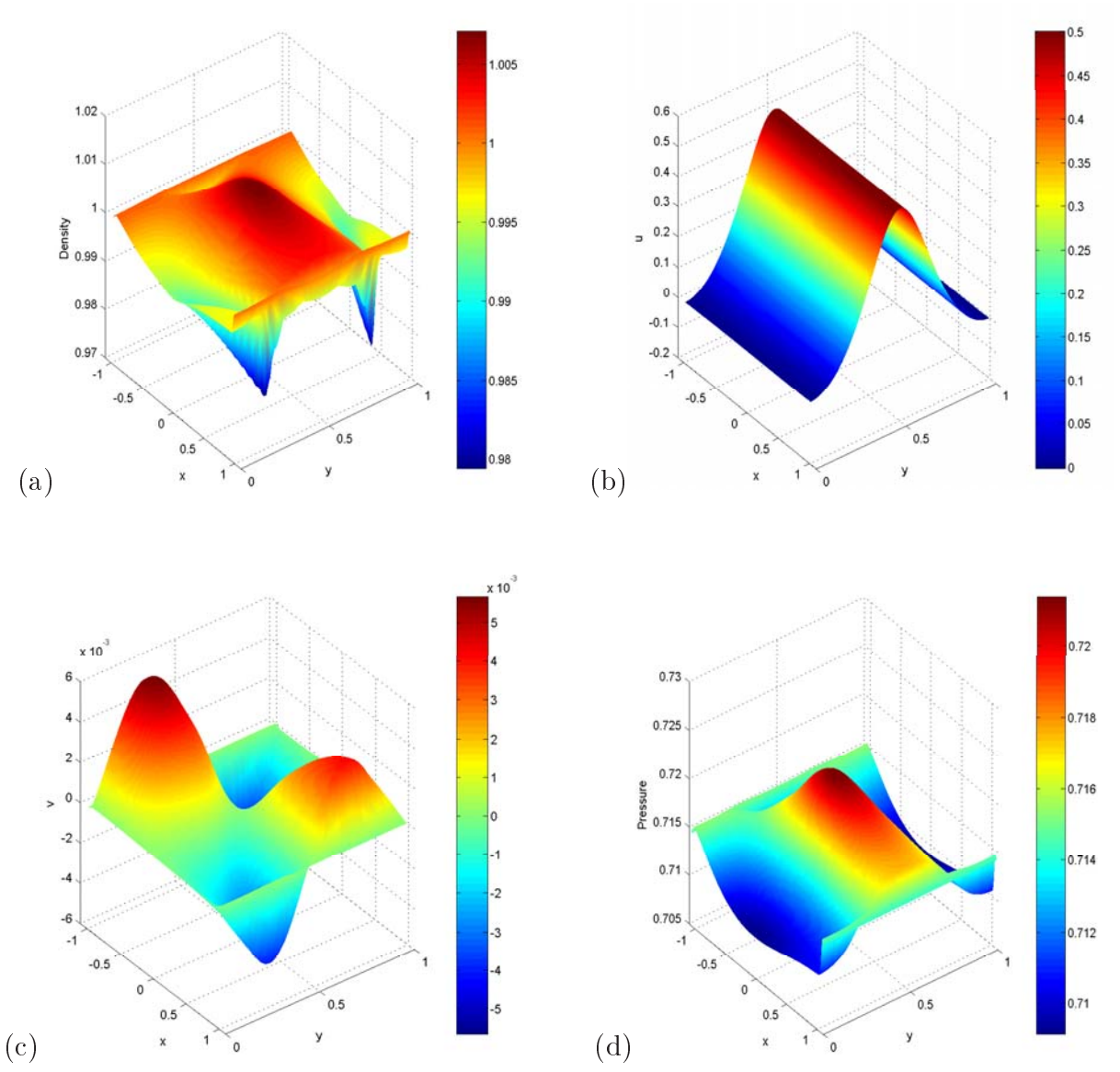


Figure 5.13: Numerical solution with PML ABC at $t = 10$. (a) Density; (b) u; (c) v; (d) Pressure.

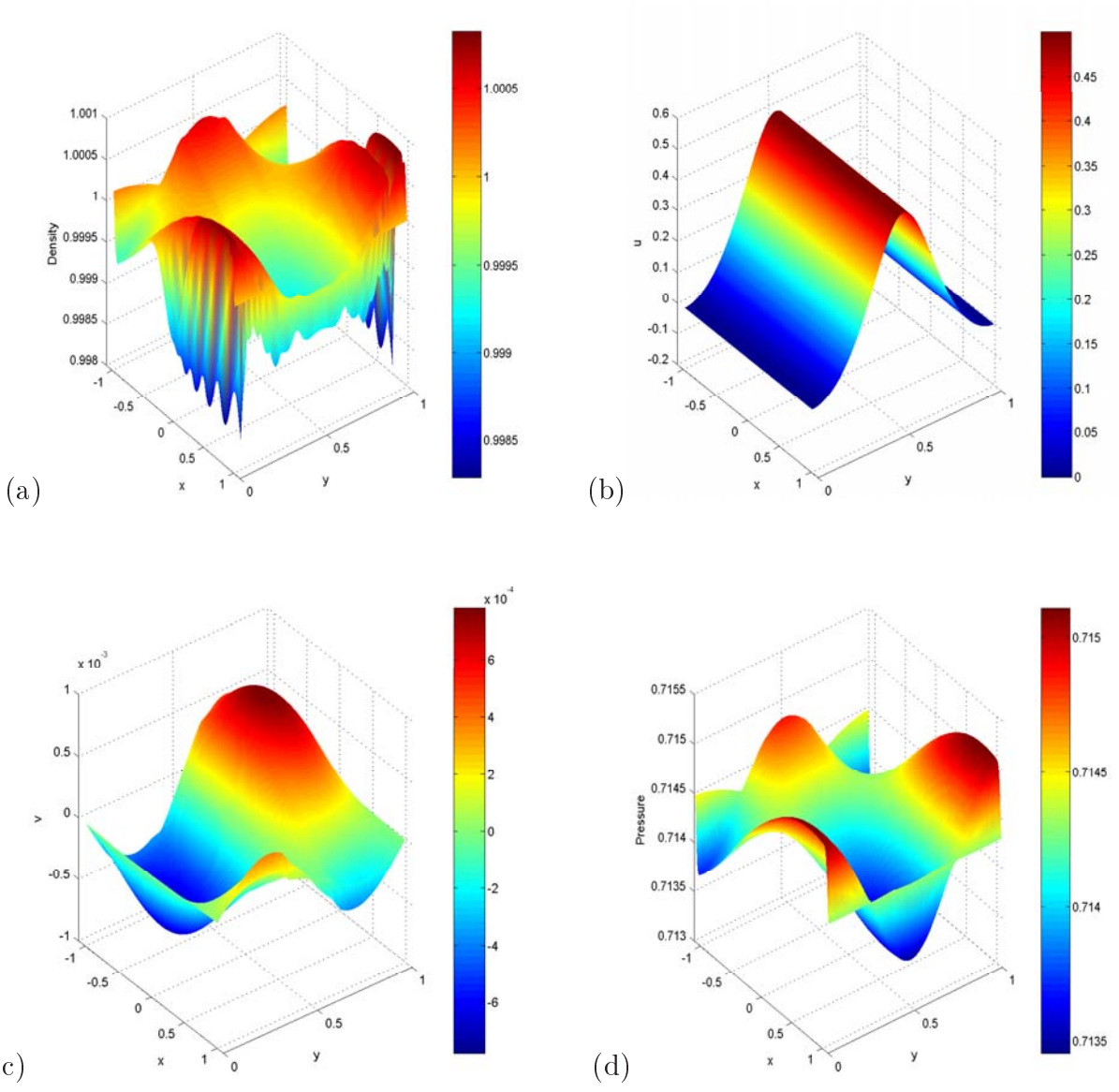


Figure 5.14: Numerical solution with PML ABC at $t = 100$. (a) Density; (b) u; (c) v; (d) Pressure.

Bibliography

[1] S. Abarbanel and D. Gottlieb, *A mathematical analysis of the PML method,* J. Comput. Phys., 134: 357-363, 1997.

[2] S. Abarbanel and D. Gottlieb, *On the construction and analysis of absorbing layers in CEM,* Appl. Numer. Math., 27: 331-340, 1998.

[3] S. Abarbanel, D. Gottlieb, and J.S. Hesthaven, *Well-posed perfectly matched layers for advective acoustics,* J. Comput. Phys., 154: 266-283, 1999.

[4] T. Amro, A. Arnold, and C. Zheng, *Numerical solution to nonlinear Klein-Gordon equations by the PML approach,* preprint 2007.

[5] X. Antoine, A. Arnold, C. Besse, M. Ehrhardt, and A. Schädle, *A Review of Transparent and Artificial Boundary Conditions Techniques for Linear and Nonlinear Schrdinger Equations,* submitted to Applied Numerical Mathematics, 2007.

[6] D. Appelö and G. Kreiss, *Evaluation of a Well-posed Perfectly Matched Layer for Computational Acoustic,* Proceedings of the HYP2002 conference, 2002, Pasadena, USA.

[7] D. Appelö, T. Hagstrom, and G. Kreiss, *Perfectly matched layers for hyperbolic systems: general formulation, well-posedness and stability,* SIAM Appl. Math. 67(1): 1-23, 2006.

[8] A. Arnold, *On Absorbing boundary conditions for quantum transport equations*, Math. Meth. Num. Anal., 28: 853-872, 1994.

[9] U.M. Ascher, S.J. Ruuth, and R.J. Spiteri, *Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations,* Appl. Numer. Math., 25(2-3): 151-167, 1997.

[10] A. Barry, J.Bielak, and R. MacCamy, *On absorbing boundary conditions for wave propagation*, J. Comput. Phys., 449-468, 1979.

[11] A. Bayliss and E. Turkel, *Radiation boundary conditions for wave-like equations*, Comm. Pure Appl. Math., 33:707-725, 1980.

[12] A. Bayliss and E. Turkel, *Far field boundary conditions for compressible flows*, J. Comput. Phys., 48 (2): 182-199, 1982.

[13] J.P. Bérenger, *A perfectly matched layer for the absorption of electromagnetic waves,* J. Comput. Phys., 114: 185-200, 1994.

[14] D. Carlson and H. Schneider, *Inertia theorems for matrices : the semidefinit case*, J. Math. Anal. Appl., 6: 430-446, 1963.

[15] W.C. Chew and W.H. Weedon, *A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates,* Microwave Opt. Technol. Lett., 7: 599-604, 1994.

[16] F. Collino and P. Monk, *The perfectly matched layer in curvilinear coordinates,* SIAM J. Sci. Comput., 19(6): 2061-2090, 1998.

[17] F. Collino and P. Monk, *Optimizing the perfectly matched layer,* Comput. Methods Appl. Mech. Engrg., 164: 157-171, 1998.

[18] T. Colonius, *Modeling artificial boundary conditions for compressible flow,* Annual Review of Fluid Mechanics, 36: 315-345, 2004.

[19] T. Colonius, S.K. Lele, and P. Morin, *Boundary conditions for direct computation of aerodynamic sound generation,* AIAA paper, 31(9): 1574-1582, 1993.

[20] R.K. Dodd, J.C. Eilbeck, J.D. Gibbon, and H.C. Morris, *Solitons in nonlinear wave equations,* Academic Press, New York, 1982.

[21] G. Doetsch, *Guide to the applications of the Laplace and z-transforms*, Van Nostrand Reinholf Company, 1971.

[22] P.J. Drazin and R.S. Johnson, *Solitons: an introduction,* Cambridge University Press, 1989.

[23] B. Engquist and L. Halpern, *Far field boundary conditions for computation over long time* , Appl. Num. Math., 4: 21-45, 1988.

[24] B. Engquist and A. Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comput., 31: 629-651, 1977.

[25] B. Engquist and A. Majda, *Radiation boundary conditions for acoustic and elastic wave calculations*, Comm. Pure Appl. Math., 32: 313-357, 1979.

[26] L.C. Evans, *Partial differential equations*, AMS, 1998.

[27] A.S. Fokas, *The generalized Dirichlet-to-Neumann map for certain nonlinear evolution PDEs,* Comm. Pure Appl. Math., 58(5): 639-670, 2005.

[28] M. Giles, *Nonreflecting boundary conditions for Euler equation calculations*, AIAA J., 28: 2050-2058, 1990.

[29] D. Givoli, *Numerical methods for problems in infinite domains*, Elsevier, Amsterdam, 1992.

[30] D. Givoli, *Nonreflecting boundary conditions*, J. Comput. Phys., 94: 1-29, 1991.

[31] D. Givoli, *High-order local non-reflecting boundary conditions: a review*, Wave Motion, 39: 319-326, 2004.

[32] D. Givoli and B. Neta, *High-order non-reflecting boundary conditions for dispersive waves*, Wave Motion, 37: 257-271, 2003.

[33] D. Givoli and I. Patlashenko, *Dirichlet-to-Neumann boundary condition for time-dependent dispersive waves in three-dimensional guides*, J. Comput. Phys., 199: 339-354, 2004.

[34] J. Goodrich and T. Hagstrom, *A comparison of two accurate boundary treatments for computational aeroacoustics*, Technical Report 97-1585, AIAA, 1999.

[35] J. Grote, *Am Rande des Unendlichen : Numerische Verfahren für unbegrenzte Gebiete*, Elemnet der Mathematik 55: 67-83, 2000.

[36] B. Gustafsson, H-O. Kreiss, and J. Oliger, *Time dependent problems and difference method*, John Wiley & Sons Inc., 1995.

[37] B. Gustafsson, H-O. Kreiss, and Sundström, *Stability theory of difference approximations for mixed initial boundary value problems*, Math. Comput., 26: 649-686, 1972.

[38] M. Goldberg and E. Tadmor, *Convenient stability criteria for difference approximations of hyperbolic initial boundary value problems II*, Math. Comput., 48: 503-520, 1987.

[39] T. Hagstrom, *On high-order radiation boundary conditions*, In B. Engquist and G. Kriegsmann, editors, IMA Volume on Computational Wave Propagation, Springer, New York, 122, 1996.

[40] T. Hagstrom, *Radiation boundary conditions for the numerical simulation of waves,* Acta Numerica, 8: 47-106, 1999.

[41] T. Hagstrom, *New results on absorbing layers and radiation boundary conditions,* Topics in Computational Wave Propagation, M. Ainsworth *et al.* eds., Springer-Verlag, 1-42, 2003.

[42] T. Hagstrom, *A new construction of perfectly matched layers for hyperbolic systems with applications to the linearized Euler equations.* G. Cohen, E.Heikkola, P. Joly and P. Neittaanmäki, eds., Proceeding of the Sixth International Conference on Mathematical and Numerical Aspects of Wave Propagation Phenomena, Springer-Verlage, 125-129, 2003.

[43] T. Hagstrom and I. Nazarov, *Absorbing layers and radiation boundary conditions for jet flow simulations,* AIAA paper 2002-2606, 2002.

[44] T. Hagstrom and I. Nazarov, *Perfectly matched layers and radiation boundary conditions for shear flow calculations,* AIAA paper 2003-3298, 2003.

[45] H.D. Han, X.N. Wu, and Z.L. Xu, *Artificial boundary method for Burgers' equation using nonlinear boundary conditions,* to appear in J. Comput. Math.

[46] E. Harrier, C. Lubich, and M. Schlichte, *Fast numerical solution of nonlinear volterra convolutional equations*, SIAM J. Sci. Stat. Comput., 6: 532-541, 1985.

[47] R.L. Higdon, *Initial boundary value problems for linear hyperbolic systems*, SIAM Review 28: 177-217, 1986.

[48] R.L. Higdon, *Absorbing boundary conditions for difference approximations to the multidimensional wave equation*, Math. Comp., 47:437459, 1986.

[49] J.S. Hesthaven, *On the analysis and construction of perfectly matched layers for the linearized Euler equations,* J. Comp. Phys. 142: 129-147, 1998.

[50] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, 1990.

[51] F.Q. Hu, *On absorbing boundary conditions for linearized Euler equations by a perfectly matched layer,* J. Comput. Phys., 129: 201-219, 1996.

[52] F.Q. Hu, *A stable, perfectly matched layer for linearized Euler equations in unsplit physical variables,* J. Comput. Phys., 173: 455-480, 2001.

[53] F.Q. Hu, *Absorbing boundary conditions,* Inter. J. Comput. Fluid Dynamics, 18(6): 513-522, 2004.

[54] F.Q. Hu, *A perfectly matched layer absorbing boundary condition for linearized Euler equations with a non-uniform flow,* J. Comput. Phys., 208: 469 - 492, 2005.

[55] F.Q. Hu, *On the construction of PML absorbing boundary condition for the non-linear Euler equations,* AIAA paper 2006-0798, 2006.

[56] F.Q. Hu and H.L. Atkins, *Eigensolutions analysis of the discontinuous Galerkin method with nonuniform grids. Part I: one space dimension,* J. Comput. Phys., 182: 516-545, 2002.

[57] F.Q. Hu, X.D. Li, and D.K. Lin, *PML absorbing boundary condition for non-linear aeroacoustics problems,* AIAA paper 2006-2521, 2006.

[58] M. Israeli and S.A. Orszag, *Approximation of radiation boundary conditions,* J. Comput. Phys., 41: 115-135, 1981.

[59] R. Kosloff and D. Kosloff, *Absorbing boundary conditions for wave propagation problems,* J. Comput. Phys., 63: 363-376, 1986.

[60] H.-O. Kreiss, *Stability theory for difference approximations of mixed initial boundary value problems*, Math. Comput., 22: 703-714, 1968.

[61] H.-O. Kreiss, *Initial boundary value problems for hyperbolic systems*, Comm. Pure and Appl. Math., 23: 277-298, 1970.

[62] H.-O. Kreiss and J. Lorenz, *Initial boundary value problems and the Navier-Stokes equations*, Academic Press Inc., 1989.

[63] M. Kuzuoglu and R. Mittra, *Frequency dependence of the constitutive parameters of causal perfectly matched absorbers,* IEEE Microwave Guided Wave Lett., 6: 447-449, 1996.

[64] R.J. Leveque, *Finite volume methods for hyperbolic problems,* Cambridge Uni. Press, 2002.

[65] M. Nakamura and T. Ozawa, *The Cauchy problem for nonlinear Klein-Gordon equations in the Sobolev spaces,* Publ. Res. Inst. Math. Sci., 37: 255-293, 2001.

[66] I.M. Navon, B. Neta, and M.Y. Hussaini, *A perfectly matched layer formulation for the nonlinear shallow water equations models: The split equation approach,* submitted to Mon. Wea. Rev., 2001.

[67] S. Osher, *Stability of difference approximations of dissipative type fo mixed initial-boundary value problems, I*, Math. Comput., 23: 335-340, 1969.

[68] L. Pareschi and G. Russo, *Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation,* J. Sci. Comp., 25: 129-155, 2005.

[69] M.M. Rai and P. Morin, *Direct numerical simulation of transition and turbulence in a spatially evolving boundary layer,* AIAA paper, 91-1607, 1991.

[70] R. D. Richtmyer and K. W. Morton, *Difference methods for initial-value problems*, Wiley-interscience, 1967.

[71] C. W. Rowley and T. Colonius, *Discretely nonreflectiing boundary conditions for linear hyperbolic systems*, J. Comput. Phys., 157(2):500-538, 2000.

[72] J.C. Strikwerda, *Finite difference schemes and partial differential equations*, Chapman and Hall, 1989.

[73] J. Szeftel, *A nonlinear approach to absorbing boundary conditions for the semilinear wave equation,* Math. Comp., 75: 565-594, 2006.

[74] J. Szeftel, *Absorbing boundary conditions for nonlinear scalar partial differential equations,* Comput. Methods Appl. Mech. Engrg., 195: 3760-3775, 2006.

[75] C.K.W. Tam, L. Auriault, and F. Cambulli, *Perfectly Matched Layer as an absorbing boundary condition for the linearized Euler equations in open and ducted domains,* J. Comput. Phys., 144: 213-234, 1998.

[76] J.W. Thomas, *Numerical partial differential equations, finite difference methods*, Springer, 1995.

[77] L.N. Trefethen and L. Halpern, *Well-posedness of one-way wave equations and absorbing boundary conditions*, Math. Comp., 176: 421435, 1986.

[78] S.V. Tsynkov, *Numerical solution of problems on unbounded domains. A review,* Appl. Numer. Math., 27: 465-532, 1998.

[79] E. Turkel and A. Yefet, *Absorbing PML boundary layers for wave-like equations,* Appl. Numer. Math., 27: 533-557, 1998.

[80] S. Weinberg, *The quantum theory of fields,* Cambridge University Press, 1995.

[81] M.A. Zahid and M.N. Guddati, *Padded continued fraction absorbing boundary conditions for dispersive waves,* Comput. Methods Appl. Mech. Engrg., 195: 3797-3819, 2006.

[82] C. Zheng, *Numerical solution to the sine-Gordon equation defined on the whole real axis,* in preprint.

[83] C. Zheng and T. Amro, *A PML absorbing boundary condition for the nonlinear Euler equations in unbounded domains,* submitted to J. Comput. Phys., 2007.

List of Figures