Well-posedness and Stability of Exact Non-reflecting Boundary Conditions

Sofia Eriksson and Jan Nordström

Original Publication:

Sofia Eriksson and Jan Nordström, Well-posedness and Stability of Exact Non-reecting Boundary Conditions, 2013, AIAA Aerospace Sciences - Fluid Sciences Event, 1-18.
http://dx.doi.org/10.2514/6.2013-2960

From the 21st AIAA Computational Fluid Dynamics Conference 24 - 27 June 2013 San Diego, California

Well-posedness and Stability of Exact Non-reflecting Boundary Conditions

Sofia Eriksson*

Department of Information Technology, Uppsala University, SE-751 05 Uppsala, Sweden

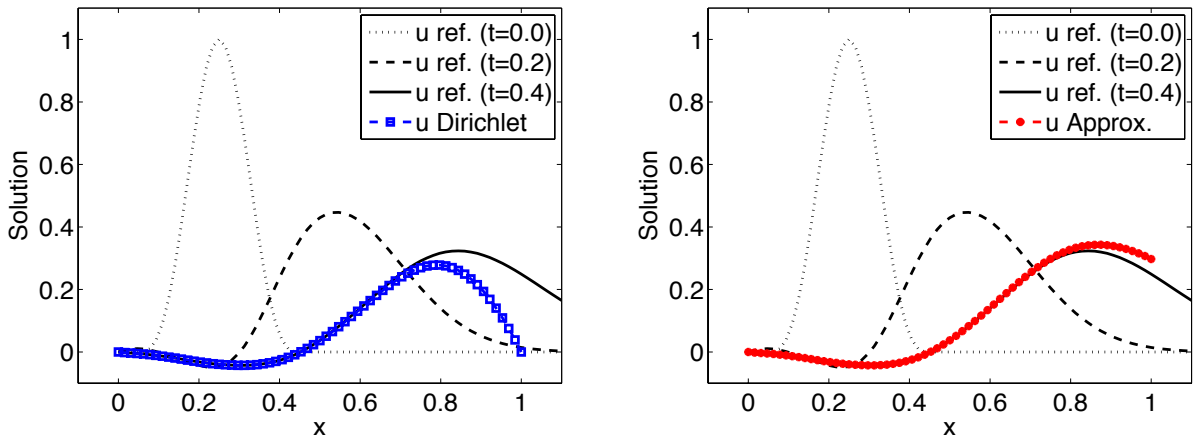
Jan Nordström†

Department of Mathematics, University of Linköping, SE-581 83 Linköping, Sweden

Exact non-reflecting boundary conditions for an incompletely parabolic system have been studied. It is shown that well-posedness is a fundamental property of the non-reflecting boundary conditions. By using summation by parts operators for the numerical approximation and a weak boundary implementation, energy stability follows automatically. The stability in combination with the high order accuracy results in a reliable, efficient and accurate method. The theory is supported by numerical simulations.

I. Introduction

In computational physics one often encounters the problem of how to limit the computational domain. For example, when simulating the flow field around an aircraft it is impossible to include the entire atmosphere. It is therefore necessary to truncate the domain at some distance away from the area of interest and introduce artificial boundary conditions (ABC). Such boundaries will generate non-physical disturbances, and in many applications it is essential that these disturbances are minimized.



(a) Solution at $t = 0.4$

(b) Solution at $t = 0.4$

Figure 1. The solution to equation (1), with initial condition given by (50). At $x = 1$ the pulse should pass out without reflections. At the right boundary either (a) a Dirichlet boundary condition or (b) a zeroth order approximate NRBC is imposed.

If the errors produced at the boundary stay localized, the boundary conditions have limited influence over the flow field and a simple boundary condition, of Dirichlet type could be used. However, this assumption

*PhD, Uppsala University, SE-751 05 Uppsala, Sweden.

†Professor, University of Linköping, SE-581 83 Linköping, Sweden.

is seldom valid, and when a wave encounters the boundary a significant portion will reflect back. This is illustrated in Figure 1(a), where the Dirichlet boundary condition pollutes the whole solution.

Apparently a better strategy is needed. In the classical paper,[7] exact boundary closures are constructed in transformed space for the wave equation. The approach is to express the solution as a superposition of waves, and eliminate the incoming waves at the boundaries. Similar techniques for deriving the non-reflecting boundary conditions (NRBC), for other types of equations, are used in.[12,16,18] Note that these conditions are exact, but formulated in transformed space.

Exact NRBC's are in most cases global in space and time, and can therefore be cumbersome to implement numerically. For that reason it is common to approximate or localize the NRBC's in space or time. In,[7] where the exact NRBC's are made local in space and time using expansions, it is shown that some approximations are well-posed, and some ill-posed. To achieve boundary conditions that give sufficiently small reflections, high order expansions are necessary, which typically yields an ill-posed problem. For low order expansions, which result in a Dirichlet-to-Neumann map, it is easier to obtain well-posedness, and the results are still clearly better than the results obtained using the Dirichlet boundary condition, see Figure 1(b).

The main drawback with the approximative NRBC's are that they ruin the increased accuracy expected from mesh refinement of the interior scheme. From Table 1 it is evident that the solution obtained using the approximate NRBC's, although it looks promising in Figure 1(b), does not converge to the correct solution as we refine the mesh. There will always be an order one error remaining in the solution.

N	Error(u)	ratio	conv. rate
16	0.01385376		
32	0.01407668	0.9842	-0.0230
64	0.01409021	0.9990	-0.0014
128	0.01409115	0.9999	-0.0001

Table 1. Results obtained using the approximative NRBC.

The area of ABC's has been the subject of massive research, see for example,[2,9,11,13,17] where the approach in[2,17] yields local boundary conditions that can be made arbitrarily accurate. When it comes to the implementation of exact NRBC's it is, for special geometries, possible to localize the boundary conditions in time while still keeping them exact. This is exemplified in[10] where computations are performed for the wave equation on a spherical domain, and in[30] where highly accurate boundary conditions are used for a flow in a cylinder. See[16] for more details on exact and approximate NRBC on special computational domains. These techniques are unfortunately not always feasible, and in[23] computations are performed for the Schrödinger equation with the exact NRBC's using convolution quadratures. For an extensive review on ABC's, see.[33] An alternative to the above mentioned methods is to introduce buffer zones outside the artificial boundary, where the governing equations are modified such that waves are damped. When these zones are constructed to be exactly non-reflecting for the continuous problem, they are called perfectly matched layers (PML), see.[1,3,19]

In this paper we follow the work in[7] to some extent, but consider a slightly different problem and most importantly; no approximations will be used. Our main interest is the theoretical aspects of the problem, i.e. the well-posedness and stability properties of exact NRBC's. The exact boundary conditions are derived in the Laplace transformed space, and thereafter transformed back for the numerical simulations. The boundary conditions are hence global in time. We use high order accurate finite difference techniques, see,[4,26,28,32] such that the error originating from the interior discretization is kept at a minimum.

The main point of this paper is that we show that the exact NRBC's result in a well-posed problem, and that this leads to energy estimates both for the continuous and the discrete formulation of the problem. We can thus, by a chain of arguments, guarantee a stable numerical procedure. The stability in combination with the high order accuracy results in a reliable, efficient and accurate method.

The paper is organized as follows. In section 2 we formulate the continuous problem. In section 3 exact non-reflecting boundary conditions are derived. In section 4 we show that the continuous problem is well-posed when using the non-reflecting boundary conditions, and that this leads to an energy estimate. The corresponding semi-discrete problem is presented in section 5. In section 6, two different approaches to choose the boundary procedure are presented, both leading to energy stability. Then, in section 7, the boundary conditions which are derived in the Laplace transformed space are transformed to physical space

using convolution quadratures. In section 8 numerical experiments are presented and conclusions are drawn in section 9.

II. The continuous problem formulation

Consider the linear 2×2 system of partial differential equations

$$\begin{aligned} U_t + AU_x - BU_{xx} &= F, \qquad x \in [x_L, x_R], \qquad t \geq 0 \\ U &= f, \qquad x \in [x_L, x_R], \qquad t = 0 \\ L_{L,R}U &= g_{L,R}, \qquad x = x_{L,R}, \qquad t \geq 0, \end{aligned} \tag{1}$$

where

$$U = \begin{bmatrix} p \\ u \end{bmatrix}, \qquad A = \begin{bmatrix} v & c \\ c & v \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 \\ 0 & \varepsilon \end{bmatrix}, \qquad v > 0.$$

$F(x,t)$ is the forcing function and $f(x)$ is the initial data. The operators L_L and L_R and the data g_L and g_R in the boundary conditions $L_{L,R}U = g_{L,R}$ are at this stage unknown. The Initial Boundary Value Problem (IBVP) (1) is incompletely parabolic and hence it has most of the properties and difficulties associated with the compressible Navier-Stokes equations. Throughout the paper we assume $v > 0$. Exactly the same analysis can be done for negative values of v.

The Laplace transformed version of (1) is

$$\begin{aligned} s\hat{U} + A\hat{U}_x - B\hat{U}_{xx} &= \hat{F} + f, \qquad x \in [x_L, x_R] \\ \hat{L}_{L,R}\hat{U} &= \hat{g}_{L,R}, \qquad x = x_{L,R}, \end{aligned} \tag{2}$$

where $s = \eta + \xi i$ is the dual variable to time, and $\hat{U} = [\ \hat{p},\ \hat{u}\ ]^T$ is defined as

$$\hat{U}(x,s) = \mathcal{L}\left\{U(x,t)\right\} = \int_0^\infty e^{-st}U(x,t)\,dt, \qquad\qquad \mathcal{L}\{U'(x,t)\} = s\hat{U}(x,s) - U(x,0).$$

To simplify the analysis, we write (2) on first order form by introducing $\hat{w} = \hat{u}_x$, which yields

$$\begin{aligned} \bar{S}\bar{U} + \bar{A}\bar{U}_x &= \bar{F}, \qquad x \in [x_L, x_R] \\ \bar{L}_{L,R}\bar{U} &= \hat{g}_{L,R}, \qquad x = x_{L,R}, \end{aligned} \tag{3}$$

where $\bar{S} = diag(s, s, 1)$ and where

$$\bar{A} = \begin{bmatrix} v & c & 0 \\ c & v & -\varepsilon \\ 0 & -1 & 0 \end{bmatrix}, \qquad \bar{U} = \begin{bmatrix} \hat{p} \\ \hat{u} \\ \hat{w} \end{bmatrix}, \qquad \bar{F} = \begin{bmatrix} \hat{F}_1 + f_1 \\ \hat{F}_2 + f_2 \\ 0 \end{bmatrix}. \tag{4}$$

The solution to (3) consists of a homogenous and a particular part, such that $\bar{U} = \bar{U}_h + \bar{U}_p$. The particular solution $\bar{U}_p$ (which depends on the data $\bar{F}$) is assumed to be known. The ansatz $\bar{U}_h = e^{\kappa x}\Psi$ leads to a generalized eigenvalue problem for $\kappa(s)$ and $\Psi(s)$ on the form

$$(\bar{S} + \kappa\bar{A})\Psi = 0. \tag{5}$$

The eigenvalue problem (5) can have non-trivial solutions $\Psi \neq 0$ if the determinant $|\bar{S} + \kappa\bar{A}|$ is zero. Written out explicitly the determinant is

$$|\bar{S} + \kappa\bar{A}| = q(\kappa, s), \qquad\qquad q(\kappa, s) = s^2 + 2sv\kappa + (v^2 - c^2 - s\varepsilon)\kappa^2 - \varepsilon v\kappa^3. \tag{6}$$

Solving $q(\kappa, s) = 0$ for the eigenvalues κ, and assuming that the three roots κ_j are distinct, gives the general homogeneous solution

$$\bar{U}_h = \sum_{j=1}^{3} \sigma_j e^{\kappa_j x}\Psi_j. \tag{7}$$

The coefficients σ_j can be determined by using the boundary conditions. This procedure is described in detail in.[14,27]

Remark: The solution $\bar{U}_h$ can be written on the form given in (7) unless $s = 0$ at the same time as $v = c$, since $q(\kappa, s)$ in (6) then has a multiple root. In the rest of the paper we assume $v \neq c$.

III. Derivation of the boundary conditions

Before the boundary conditions are constructed it is essential to know how many that are needed at each boundary. It is shown in[31] that for each negative $\mathrm{Re}(\kappa)$ we need one condition at the left boundary, and for each positive $\mathrm{Re}(\kappa)$ we need one condition at the right boundary. The number of roots with negative and positive real parts, respectively, is given by

Proposition III.1. *Consider the roots of $q(\kappa, s) = 0$ in (6). For $v > 0$ and s such that $Re(s) > 0$, two of the κ's have negative real part and one of the κ's has positive real part.*

Proof. Assume that κ passes the imaginary axis, i.e. that $\kappa = \beta i$. Inserting this into equation (6) and using that $s = \eta + \xi i$ yields

$$c^2\beta^2 + \varepsilon\eta\beta^2 + \eta^2 - (\xi + v\beta)^2 + (2\eta + \varepsilon\beta^2)(\xi + v\beta)i = 0. \tag{8}$$

The imaginary part of (8) is zero if either $\xi + v\beta = 0$ or $2\eta + \varepsilon\beta^2 = 0$. In both of these cases, it is required that either $\eta < 0$ or that $\eta = \xi = 0$ to cancel the real part. That is, as long as the real part of s is positive ($\eta > 0$), no purely imaginary κ can exist and hence the real part of the κ's can not change sign. Dividing $q(\kappa, s)$ in (6) by $-\varepsilon v$ yields

$$\begin{gathered}\tilde{q}(\kappa, s) = \kappa^3 - \underbrace{\frac{(v^2 - c^2 - s\varepsilon)}{\varepsilon v}}_{r_2}\kappa^2 + \underbrace{\frac{-2s}{\varepsilon}}_{r_1}\kappa - \underbrace{\frac{s^2}{\varepsilon v}}_{r_0} = (\kappa - \kappa_1)(\kappa - \kappa_2)(\kappa - \kappa_3) \\ r_2 = \kappa_1 + \kappa_2 + \kappa_3, \qquad r_1 = \kappa_1\kappa_2 + \kappa_1\kappa_3 + \kappa_2\kappa_3, \qquad r_0 = \kappa_1\kappa_2\kappa_3,\end{gathered} \tag{9}$$

and by assuming s real and large, we get $r_0 > 0$, $r_1 < 0$ and $r_2 < 0$. According to Descartes' rule of signs,[29] the polynomial $\tilde{q}(\kappa, s)$ has exactly one positive root for these values of r_0, r_1 and r_2. □

Thus two boundary conditions are needed at the left boundary and one boundary condition is needed at the right boundary. Without loss of generality, let $\mathrm{Re}(\kappa_1) < 0$, $\mathrm{Re}(\kappa_2) < 0$ and $\mathrm{Re}(\kappa_3) > 0$.

A. Non-reflecting boundary conditions

One approach when constructing non-reflecting boundary conditions is to prohibit the solution outside the artificial boundary from growing, i.e. by demanding that $\bar{U}_h(x) \to 0$ as $x \to \pm\infty$, see.[33] This is accomplished by canceling the coefficients σ_j in (7) corresponding to the growing modes at each boundary.

Remark: Compare with the hyperbolic version of (1), where the characteristics of $U(x, t)$ travel with constant wave speed a_j. In this case the eigenvalues of the Laplace transformed solution have the form $\kappa_j = -s/a_j$ and the eigenvectors Ψ_j are independent of s, such that

$$U(x, t) = \sum_j h_j(t - x/a_j)\Psi_j, \qquad \hat{U}(x, s) = \sum_j \hat{h}_j(s)e^{-xs/a_j}\Psi_j.$$

Thus a positive wave speed a_j, which means that the eigensolution Ψ_j is right-going, implies that $\mathrm{Re}(\kappa_j) = -\mathrm{Re}(s)/a_j$ is negative. Likewise, if $\mathrm{Re}(\kappa_j) > 0$, the eigenfunction Ψ_j is left-going. For a hyperbolic problem, providing zero data directly to the ingoing variables means that the outgoing waves can pass through the boundary freely, without reflections. Analogously, in (7) we should at each boundary cancel the modes that are growing outwards.

Recall that the real parts of κ_1 and κ_2 are negative and the real part of κ_3 is positive for $\mathrm{Re}(s) > 0$. Our aim is to construct boundary conditions for the left boundary that force σ_1 and σ_2 to zero, and a boundary

condition for the right boundary that forces σ_3 to zero. With access to an eigenvalue κ_i we compute the eigenvector Ψ_i and the corresponding orthogonal vector Φ_i

$$\Psi_i = \begin{bmatrix} -c \\ (s+v\kappa_i)/\kappa_i \\ s+vk_i \end{bmatrix}, \qquad \Phi_i = \begin{bmatrix} \varepsilon(v\kappa_j+s)(v\kappa_k+s)/sc \\ \varepsilon v\kappa_j\kappa_k/s \\ \varepsilon \end{bmatrix}, \tag{10}$$

where κ_j and κ_k are the remaining two roots (if $\kappa_i = \kappa_1$ then $\kappa_{j,k} = \kappa_{2,3}$). The vector Φ_l is orthogonal to Ψ_i for $i \neq l$, such that

$$\Phi_i^T\Psi_i = \varepsilon v(\kappa_i-\kappa_j)(\kappa_i-\kappa_k)/\kappa_i, \qquad \Phi_j^T\Psi_i = 0, \qquad \Phi_k^T\Psi_i = 0. \tag{11}$$

Using (7) and (11) we see that the boundary condition $\Phi_i^T\bar{U}_h = 0$ is equivalent to $\sigma_i e^{\kappa_i x}\Phi_i^T\Psi_i = 0$, which forces σ_i to zero. This gives the exact non-reflecting boundary conditions

$$x = x_L: \quad \begin{cases} \Phi_1^T\bar{U}_h = 0 \\ \Phi_2^T\bar{U}_h = 0 \end{cases}, \qquad x = x_R: \quad \Phi_3^T\bar{U}_h = 0. \tag{12}$$

The boundary conditions (12) are $\bar{L}_L\bar{U}_h = 0$ and $\bar{L}_R\bar{U}_h = 0$, where

$$\bar{L}_L = [\Phi_1,\ \Phi_2]^T, \qquad \bar{L}_R = \Phi_3^T. \tag{13}$$

Thus we can identify

$$\bar{L}_{L,R}\bar{U} = \bar{L}_{L,R}(\bar{U}_h + \bar{U}_p) = \bar{L}_{L,R}\bar{U}_p \qquad \Longrightarrow \qquad \hat{g}_{L,R} = \bar{L}_{L,R}\bar{U}_p.$$

Finding the data $\hat{g}_{L,R}$ can be difficult. Common choices are to assume that $\bar{U}_p$ is constant or zero. To take the possibility of non-exact data into account, assume that the boundary data has been chosen such that $\hat{g}_{L,R} = \bar{L}_{L,R}\bar{U}_p + g'_{L,R}$. Then, in practice, the boundary conditions imposed are

$$x = x_L: \quad \bar{L}_L\bar{U}_h = g'_L, \qquad x = x_R: \quad \bar{L}_R\bar{U}_h = g'_R, \tag{14}$$

where $g'_{L,R}$ should be some perturbation close to (or preferably equal to) zero.

Remark: The particular solution $\bar{U}_p$ depends on $\bar{F}$, which in turn depends on the forcing function F and the initial function f in (1). Often these functions are defined so that they have compact support, which implies that $\bar{U}_p = 0$ at the boundaries and that $\hat{g}_{L,R} = 0$. One reasonable exception is a constant non-zero background flow.

IV. Well-posedness of the IBVP in the GKS sense

The problem (1) is well-posed in the GKS[a] sense if no solutions $U(x,t)$ that grow exponentially in time exist, see.[6,14,15,25,27] (A more generous definition of well-posedness, that opens up for a wider range of problems, is to accept bounded growth of the solution. In this paper we limit ourselves to zero growth.)

Remark: A problem is well-posed (Hadamard's well-posedness) if: i) A solution exists, ii) The solution is unique, iii) The solution depends continuously on provided data. Existence is guaranteed by using the right number of boundary conditions and uniqueness follows from iii). We will focus on the third requirement, which is equivalent to limit the growth of the solution, see.[14]

Consider the homogeneous solution (7). By defining

$$\Psi = [\Psi_1, \Psi_2, \Psi_3], \qquad K(x) = diag(e^{\kappa_1 x},\ e^{\kappa_2 x},\ e^{\kappa_3 x}), \qquad \sigma = [\sigma_1, \sigma_2, \sigma_3]^T,$$

[a]GKS refers to the classical paper[15] by Gustafsson, Kreiss and Sundström.

we can write $\bar{U}_h = \Psi K \sigma$. Next, the boundary conditions in (14) are applied, yielding

$$E(s)\sigma = g', \qquad E(s) = \begin{bmatrix} \bar{L}_L \Psi K(x_L) \\ \bar{L}_R \Psi K(x_R) \end{bmatrix} = \begin{bmatrix} e^{\kappa_1 x_L}\Phi_1^T\Psi_1 & e^{\kappa_2 x_L}\Phi_1^T\Psi_2 & e^{\kappa_3 x_L}\Phi_1^T\Psi_3 \\ e^{\kappa_1 x_L}\Phi_2^T\Psi_1 & e^{\kappa_2 x_L}\Phi_2^T\Psi_2 & e^{\kappa_3 x_L}\Phi_2^T\Psi_3 \\ e^{\kappa_1 x_R}\Phi_3^T\Psi_1 & e^{\kappa_2 x_R}\Phi_3^T\Psi_2 & e^{\kappa_3 x_R}\Phi_3^T\Psi_3 \end{bmatrix},$$

where $g' = [(g'_L)^T, (g'_R)^T]^T$. Each row of the system above corresponds to one boundary condition, and for general boundary conditions the matrix $E(s)$ is full. If $E(s)$ is non-singular we can solve for σ and obtain a unique solution $\bar{U} = \bar{U}_p + \Psi K E(s)^{-1} g'$. Recalling that the first two entries of $\bar{U}$ are denoted $\hat{U}$, we can formally transform back to the time domain, as

$$U(x,t) = \mathcal{L}^{-1}\{\hat{U}\} = e^{\eta_0 t}\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{U}(x, \eta_0 + i\xi) e^{i\xi t} d\xi\right)$$

where $E(s)$ must be non-singular for $\eta > \eta_0$. The problem is well-posed in our restrictive GKS sense if $\eta_0 \leq 0$, (for convergence to steady-state $\eta_0 < 0$ is necessary).

Proposition IV.1. *Consider the ordinary differential equation* (3) *with boundary operators* (13)*. The corresponding matrix $E(s)$ is non-singular for Re$(s) \geq 0$ (if $0 < v \neq c$), and hence the problem* (1) *is well-posed.*

Proof. Using that $\Phi_j^T \Psi_i = 0$ for $i \neq j$ leads to

$$E(s) = \begin{bmatrix} e^{\kappa_1 x_L}\Phi_1^T\Psi_1 & 0 & 0 \\ 0 & e^{\kappa_2 x_L}\Phi_2^T\Psi_2 & 0 \\ 0 & 0 & e^{\kappa_3 x_R}\Phi_3^T\Psi_3 \end{bmatrix}.$$

From (11) we know that $\Phi_i^T\Psi_i = \varepsilon v(\kappa_i - \kappa_j)(\kappa_i - \kappa_k)/\kappa_i$ and thereby the three entries of $E(s)$ are non-zero if the roots κ_i, κ_j, κ_k are distinct. In Appendix A it is shown that there are no multiple roots for Re$(s) \geq 0$, unless $s = 0$. This special case is treated separately, and it can be shown that $\lim_{s\to 0} \Phi_j^T \Psi_j \neq 0$ as long as $v \neq c$. Consequently $|E(s)| \neq 0$ for all Re$(s) \geq 0$ when $v \neq c$. □

A. Well-posedness in the energy sense

Proposition IV.1 above shows that the exact non-reflecting boundary conditions yield well-posedness (in the GKS sense). Next we show that the non-reflecting boundary conditions also leads to an energy estimate.

Equation (2) is multiplied by the conjugate transpose of $\hat{U}$ (denoted $\hat{U}^*$) from the left and integrated with respect to x. Adding the complex conjugate of the resulting relation to itself, and using that $s = \eta + \xi i$, we get

$$2\eta \int_{x_L}^{x_R} \hat{U}^*\hat{U} dx + 2\int_{x_L}^{x_R} \hat{U}_x^* B \hat{U}_x dx = BT_L + BT_R \tag{15}$$

where

$$BT_L = \hat{U}^* A \hat{U} - \hat{U}^* B \hat{U}_x - \hat{U}_x^* B \hat{U}\Big|_{x_L}, \qquad BT_R = -\hat{U}^* A \hat{U} + \hat{U}^* B \hat{U}_x + \hat{U}_x^* B \hat{U}\Big|_{x_R}. \tag{16}$$

Note that the forcing term $\hat{F} + f$ is omitted since it does not affect well-posedness.[14] We know from the previous analysis of $E(s)$ that the operators in (13) give a well-posed problem. However, if the boundary conditions can be imposed such that the boundary terms BT_L and BT_R are non-positive we also obtain an energy estimate, which will lead directly to stability for the discrete problem.

Since we have derived the boundary conditions for the first order form in (3) we rewrite (16) on the equivalent form

$$BT_L = \bar{U}^* \widetilde{A} \bar{U}\Big|_{x_L}, \qquad BT_R = -\bar{U}^* \widetilde{A} \bar{U}\Big|_{x_R}, \qquad \widetilde{A} = \begin{bmatrix} v & c & 0 \\ c & v & -\varepsilon \\ 0 & -\varepsilon & 0 \end{bmatrix}. \tag{17}$$

Remark: Well-posedness in the GKS sense considers the homogenous solution, $\bar{U}_h = \sum_j \sigma_j e^{\kappa_j x} \Psi_j$. Computing the energy estimate for the homogenous solution instead of the total solution only involves the forcing term (which is disregarded) and hence the boundary terms in (17) holds for $\bar{U}_h$ as well as for $\bar{U}$.

Proposition IV.2. *The left boundary term in* (17) *is non-positive, i.e.* $BT_L \leq 0$.

Proof. The left boundary conditions in (12) force σ_1 and σ_2 to zero which yields the solution $\bar{U}_h = \sigma_3 e^{\kappa_3 x}\Psi_3$. Inserting this into BT_L in (17) we obtain

$$BT_L = |\sigma_3 e^{\kappa_3 x_L}|^2 \mathcal{A}_L$$

where it is possible to show, see,[8] that

$$\mathcal{A}_L = \Psi_3^* \widetilde{A} \Psi_3 \leq 0. \tag{18}$$

□

Proposition IV.3. *The right boundary term in* (17) *is non-positive, i.e.* $BT_R \leq 0$.

Proof. The right boundary condition in (12) yields the solution $\bar{U}_h = \sigma_1 e^{\kappa_1 x}\Psi_1 + \sigma_2 e^{\kappa_2 x}\Psi_2$. Inserting this into BT_R in (17) we obtain

$$BT_R = -\begin{bmatrix} \sigma_1 e^{\kappa_1 x_R} \\ \sigma_2 e^{\kappa_2 x_R} \end{bmatrix}^* \mathcal{A}_R \begin{bmatrix} \sigma_1 e^{\kappa_1 x_R} \\ \sigma_2 e^{\kappa_2 x_R} \end{bmatrix},$$

where it is possible to show, see,[8] that

$$\mathcal{A}_R = \begin{bmatrix} \Psi_1^* \widetilde{A} \Psi_1 & \Psi_1^* \widetilde{A} \Psi_2 \\ \Psi_2^* \widetilde{A} \Psi_1 & \Psi_2^* \widetilde{A} \Psi_2 \end{bmatrix} \geq 0. \tag{19}$$

□

Since the boundary terms BT_L and BT_R are negative the right hand side of (15) is bounded, which leads to $\eta \leq 0$ and an energy estimate.

Remark: In Proposition IV.2 and Proposition IV.3 we have assumed that the provided data is exact, such that $\sigma_{1,2} = 0$ at the left boundary or $\sigma_3 = 0$ at the right boundary. Later in Section VI we will also include the possibility of having non-zero (incorrect) boundary data and show that the problem is in fact strongly well-posed.

V. The semi-discrete problem formulation

The boundary operators in (13) can be written

$$\bar{L}_L = \begin{bmatrix} \alpha_1 & \beta_1 & \varepsilon \\ \alpha_2 & \beta_2 & \varepsilon \end{bmatrix}, \qquad \hat{g}_L = \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix}, \qquad \bar{L}_R = \begin{bmatrix} \alpha_3 & \beta_3 & \varepsilon \end{bmatrix}, \qquad \hat{g}_R = \begin{bmatrix} \hat{g}_3 \end{bmatrix}, \tag{20}$$

where α_j, β_j depend on s and $\kappa_j(s)$. The structure of the complementing vectors in (10) gives

$$\alpha_i = \frac{-sc}{s + v\kappa_i}, \qquad \beta_i = \frac{s}{\kappa_i}. \tag{21}$$

The boundary conditions can also be rewritten such that they are appropriate for the problem (2), as

$$\begin{aligned} \hat{L}_L\hat{U} = H_L\hat{U} + G_L\hat{U}_x = \hat{g}_L \quad &\Longleftrightarrow \quad \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} \begin{bmatrix} \hat{p} \\ \hat{u} \end{bmatrix} + \begin{bmatrix} 0 & \varepsilon \\ 0 & \varepsilon \end{bmatrix} \begin{bmatrix} \hat{p}_x \\ \hat{u}_x \end{bmatrix} = \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix} \\ \hat{L}_R\hat{U} = H_R\hat{U} + G_R\hat{U}_x = \hat{g}_R \quad &\Longleftrightarrow \quad \begin{bmatrix} \alpha_3 & \beta_3 \end{bmatrix} \begin{bmatrix} \hat{p} \\ \hat{u} \end{bmatrix} + \begin{bmatrix} 0 & \varepsilon \end{bmatrix} \begin{bmatrix} \hat{p}_x \\ \hat{u}_x \end{bmatrix} = \begin{bmatrix} \hat{g}_3 \end{bmatrix}. \end{aligned} \tag{22}$$

A. The numerical scheme

The domain $x \in [x_L, x_R]$ is discretized in space using $N+1$ equidistant grid points, as $x_i = x_L + (x_R - x_L)i/N$, where $i = 0, 1, \ldots, N$. The solution U is represented by the discrete solution vector V such that

$$V = [V_0^T, V_1^T, \ldots, V_N^T]^T, \qquad V_i(t) \approx U(x_i, t).$$

The semi-discrete scheme representing the IBVP in (1) is written

$$\begin{aligned} V_t + (D \otimes A)V - (D^2 \otimes B)V = F + ((\Sigma_0 * V)(t) - \Gamma_0) + ((\Sigma_N * V)(t) - \Gamma_N)\,, \\ V(0) = f, \end{aligned} \tag{23}$$

where the symbol $\otimes$ refers to the Kronecker product. The boundary conditions (22) are imposed weakly in (23) using the Simultaneous Approximation Term (SAT) technique, by the penalties $((\Sigma_{0,N} * V)(t) - \Gamma_{0,N}(t))$ which are yet unknown but will be derived in the Laplace transformed domain. Further, the difference operator D (which mimics $\partial/\partial x$) is on so called Summation-By-Parts (SBP) form, and hence the following holds

$$D = P^{-1}Q, \qquad Q + Q^T = e_N e_N^T - e_0 e_0^T, \qquad P = P^T > 0, \tag{24}$$

where $e_0 = [1, 0, \ldots, 0]^T$ and $e_N = [0, \ldots, 0, 1]^T$. The second derivative $\partial^2/\partial x^2$ is approximated by the wide operator D^2. For a read-up on SBP and SAT, see[5,24] and references therein. Note that we use the same notation for F, f both in the continuous and the discrete setting.

By Laplace transforming (23) the discrete representation of (2) is obtained, as

$$s\hat{V} + (D \otimes A)\hat{V} - (D^2 \otimes B)\hat{V} = \hat{F} + f + \left(\hat{\Sigma}_0\hat{V} - \hat{\Gamma}_0\right) + \left(\hat{\Sigma}_N\hat{V} - \hat{\Gamma}_N\right), \tag{25}$$

where $\hat{V}(s) = \mathcal{L}\{V(t)\}$ and where $\hat{\Sigma}_{0,N}$, $\hat{\Gamma}_{0,N}$ remains to be determined. As in the continuous case we simplify by omitting the forcing function $\hat{F} + f$. We multiply (25) by $\hat{V}^*\bar{P}$ from the left, where $\bar{P} = P \otimes I_2$, and add the conjugate transpose of the equation to itself. Thereafter using the SBP-properties in (24) yields

$$2\eta\hat{V}^*\bar{P}\hat{V} + 2(\bar{D}\hat{V})^*(P \otimes B)\bar{D}\hat{V} = BT_L^D + BT_R^D, \tag{26}$$

where $\bar{D} = D \otimes I_2$ and where

$$\begin{aligned} BT_L^D = \quad & \hat{V}_0^* A\hat{V}_0 \; - \hat{V}_0^* B(\bar{D}\hat{V})_0 \; - (\bar{D}\hat{V})_0^* B\hat{V}_0 \; + \hat{V}^*\bar{P}(\hat{\Sigma}_0\hat{V} \; - \hat{\Gamma}_0) \; + (\hat{\Sigma}_0\hat{V} \; - \hat{\Gamma}_0)^*\bar{P}\hat{V} \\ BT_R^D = {} & -\hat{V}_N^* A\hat{V}_N + \hat{V}_N^* B(\bar{D}\hat{V})_N + (\bar{D}\hat{V})_N^* B\hat{V}_N + \hat{V}^*\bar{P}(\hat{\Sigma}_N\hat{V} - \hat{\Gamma}_N) + (\hat{\Sigma}_N\hat{V} - \hat{\Gamma}_N)^*\bar{P}\hat{V}. \end{aligned} \tag{27}$$

Note the similarity between the semi-discrete relation (26) and the continuous one in (15).

The matrices $\hat{\Sigma}_{0,N}$ and the vectors $\hat{\Gamma}_{0,N}$ depend on how the boundary conditions are imposed. We use the following ansätze for the penalty terms

$$\begin{aligned} \hat{\Sigma}_0\hat{V} - \; \hat{\Gamma}_0 &= (P^{-1}e_0 \; \otimes \tau_0 \; + P^{-1}D^T e_0 \; \otimes \sigma_0 \;)(H_L\hat{V}_0 \; + \; G_L(\bar{D}\hat{V})_0 - \hat{g}_L) \\ \hat{\Sigma}_N\hat{V} - \hat{\Gamma}_N &= (P^{-1}e_N \otimes \tau_N + P^{-1}D^T e_N \otimes \sigma_N)(H_R\hat{V}_N + G_R(\bar{D}\hat{V})_N - \hat{g}_R), \end{aligned} \tag{28}$$

where all dependence of boundary data sits in $\hat{\Gamma}_{0,N}$, such that $\hat{\Gamma}_{0,N} = 0$ if $\hat{g}_{L,R} = 0$. The boundary operators $H_{L,R}$, $G_{L,R}$ are given in (22) and the penalty parameters τ_0 and σ_0 are 2×2 matrices and τ_N and σ_N are 2×1 vectors. The relations in (28) lead to

$$\begin{aligned} \hat{V}^*\bar{P}(\hat{\Sigma}_0 \; \hat{V} - \hat{\Gamma}_0) &= (\hat{V}_0^*\tau_0 \; + (\bar{D}\hat{V})_0^*\sigma_0 \;)(H_L\hat{V}_0 \; + \; G_L(\bar{D}\hat{V})_0 \; - \hat{g}_L) \\ \hat{V}^*\bar{P}(\hat{\Sigma}_N\hat{V} - \hat{\Gamma}_N) &= (\hat{V}_N^*\tau_N + (\bar{D}\hat{V})_N^*\sigma_N)(H_R\hat{V}_N + G_R(\bar{D}\hat{V})_N - \hat{g}_R). \end{aligned} \tag{29}$$

Inserting the expressions (29) into (27), the boundary terms can be written as

$$\begin{aligned} BT_L^D = {} & \begin{bmatrix} \hat{V}_0 \\ (\bar{D}\hat{V})_0 \end{bmatrix}^* \begin{bmatrix} A + \tau_0 H_L + (\tau_0 H_L)^* & -B + \tau_0 G_L + (\sigma_0 H_L)^* \\ -B + \sigma_0 H_L + (\tau_0 G_L)^* & \sigma_0 G_L + (\sigma_0 G_L)^* \end{bmatrix} \begin{bmatrix} \hat{V}_0 \\ (\bar{D}\hat{V})_0 \end{bmatrix} \\ & - \begin{bmatrix} \hat{V}_0 \\ (\bar{D}\hat{V})_0 \end{bmatrix}^* \begin{bmatrix} \tau_0 \\ \sigma_0 \end{bmatrix} \hat{g}_L - \left(\begin{bmatrix} \hat{V}_0 \\ (\bar{D}\hat{V})_0 \end{bmatrix}^* \begin{bmatrix} \tau_0 \\ \sigma_0 \end{bmatrix} \hat{g}_L \right)^* \end{aligned} \tag{30}$$

and

$$
\begin{aligned}
BT_R^D = &\begin{bmatrix} \hat{V}_N \\ (\bar{D}\hat{V})_N \end{bmatrix}^* \begin{bmatrix} -A+\tau_N H_R+(\tau_N H_R)^* & B+\tau_N G_R+(\sigma_N H_R)^* \\ B+\sigma_N H_R+(\tau_N G_R)^* & \sigma_N G_R+(\sigma_N G_R)^* \end{bmatrix} \begin{bmatrix} \hat{V}_N \\ (\bar{D}\hat{V})_N \end{bmatrix} \\
&- \begin{bmatrix} \hat{V}_N \\ (\bar{D}\hat{V})_N \end{bmatrix}^* \begin{bmatrix} \tau_N \\ \sigma_N \end{bmatrix} \hat{g}_R - \left(\begin{bmatrix} \hat{V}_N \\ (\bar{D}\hat{V})_N \end{bmatrix}^* \begin{bmatrix} \tau_N \\ \sigma_N \end{bmatrix} \hat{g}_R \right)^*,
\end{aligned}
\tag{31}
$$

respectively.

Similarly to the definition of well-posedness for the continuous problem, a numerical scheme is energy stable if the growth of the solution is bounded. As in the continuous case we limit ourselves to zero growth, which means that $\eta \leq 0$ in (26) is needed. Hence, to prove stability, we must show that the boundary terms in (30) and (31) are non-positive for zero data. In the next section, we present two distinctly different ways of choosing the penalty parameters $\tau_{0,N}$ and $\sigma_{0,N}$ such that $BT_{L,R}^D \leq 0$.

VI. Energy estimates in Laplace space

The stability requirements alone do not determine the penalty parameters $\tau_{0,N}$ and $\sigma_{0,N}$ in (28) uniquely. We will here present two different possible choices (here referred to as "replacing the indefinite terms" and "replacing the ingoing variables"), both guaranteeing a stable numerical scheme. In both cases the strategy is to first reformulate the continuous boundary terms $BT_{L,R}$ using the boundary conditions, and then to choose the penalty parameters such that the discrete boundary terms $BT_{L,R}^D$ mimic the continuous ones.

A. Replacing the indefinite terms

1. *The continuous formulation*

Using the boundary conditions (22), which are related to the system (2), it is possible to replace the indefinite terms found in the continuous boundary terms (16), and rewrite them as

$$
\begin{aligned}
BT_L = &\quad \hat{U}^* \underbrace{(A+\widetilde{H}_L+\widetilde{H}_L^*)}_{M_L} \hat{U} - \hat{U}^* \tilde{g}_L - \tilde{g}_L^* \hat{U} \Big|_{x_L} \\
BT_R = &-\hat{U}^* \underbrace{(A+\widetilde{H}_R+\widetilde{H}_R^*)}_{M_R} \hat{U} + \hat{U}^* \tilde{g}_R + \tilde{g}_R^* \hat{U} \Big|_{x_R}.
\end{aligned}
\tag{32}
$$

where $\widetilde{H}_L = S_L H_L$, $\tilde{g}_L = S_L \hat{g}_L$ and $\widetilde{H}_R = S_R H_R$, $\tilde{g}_R = S_R \hat{g}_R$. The scaling matrices S_L and S_R have the form

$$
S_L = \begin{bmatrix} a & -a \\ b & 1-b \end{bmatrix}, \qquad S_R = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{33}
$$

where $a \neq 0$ and b are arbitrary. For an energy estimate we need $M_L = A + \widetilde{H}_L + \widetilde{H}_L^*$ in (32) to be negative semi-definite and $M_R = A + \widetilde{H}_R + \widetilde{H}_R^*$ to be positive semi-definite.

Proposition VI.1. *The constants a and b in S_L in* (33) *can always be chosen such that $M_L = A + \widetilde{H}_L + \widetilde{H}_L^*$ in* (32) *is negative semi-definite.*

Proposition VI.2. *The matrix $M_R = A + \widetilde{H}_R + \widetilde{H}_R^*$ in* (32) *is positive semi-definite.*

The proofs of Proposition (VI.1) and Proposition (VI.2) are omitted, but can be found in.[8]

2. *Choice of penalty parameters for the discrete formulation*

Proposition VI.3. *Choosing the left penalty parameters as $\tau_0 = S_L$, where S_L is given in* (33)*, and σ_0 being a 2×2 zero matrix $\mathbf{0}_2$, yields a stable numerical scheme, given that Proposition VI.1 holds and under the assumption that the right boundary terms are bounded as well.*

Proof. Inserting $\tau_0 = S_L$ and $\sigma_0 = \mathbf{0}_2$ into (30), the left discrete boundary term becomes

$$BT_L^D = \hat{V}_0^*(A + \widetilde{H}_L + \widetilde{H}_L^*)\hat{V}_0 - \hat{V}_0^*\tilde{g}_L - \tilde{g}_L^*\hat{V}_0, \tag{34}$$

and according to Proposition VI.1 $M_L = A + \widetilde{H}_L + \widetilde{H}_L^*$ can be designed to be negative semi-definite. □

Proposition VI.4. *Choosing the right penalty parameters as $\tau_N = -S_R$, where S_R is given in* (33)*, and $\sigma_N = [0,\ 0]^T$, yields a stable numerical scheme, under the assumption that the left boundary terms are bounded as well.*

Proof. Inserting $\tau_N = -S_R$ and $\sigma_N = [0,\ 0]^T$ into (31), the right discrete boundary term becomes

$$BT_R^D = -\hat{V}_N^*(A + \widetilde{H}_R + \widetilde{H}_R^*)\hat{V}_N + \hat{V}_N^*\tilde{g}_R + \tilde{g}_R^*\hat{V}_N \tag{35}$$

and according to Proposition VI.2 $M_R = A + \widetilde{H}_R + \widetilde{H}_R^*$ is positive semi-definite. □

Remark: Note that when using the penalty parameters as specified in Proposition VI.3 and VI.4, the discrete boundary terms (34) and (35) mimics the continuous ones perfectly, c.f. equation (32).

B. Replacing the ingoing variables

1. *The continuous formulation*

Using the boundary conditions (20), which are related to the system (3), it is possible to replace the ingoing variables found in the continuous boundary terms (17).

Consider the matrix $\widetilde{A}$ in (17), and assume that we have found a rotation such that $\widetilde{A} = X\Lambda X^T$, where Λ is diagonal. Note that the elements of Λ are not necessarily the eigenvalues of $\widetilde{A}$, and that the vectors in X may then not be orthogonal. According to Sylvester's law of inertia, the matrices $\widetilde{A}$ and Λ will always have the same number of positive/negative eigenvalues for a non-singular X. The matrix Λ has two positive entries and one negative entry for $v > 0$, and is sorted as $\Lambda = diag(\Lambda_+, \Lambda_-)$. The vectors are divided correspondingly, $X = [x_+, x_-]$. Further, we introduce scaling matrices $J_{L,R}$ and $R_{L,R}$ as

$$\bar{L}_L = J_L(x_+^T + R_L x_-^T), \qquad \bar{L}_R = J_R(x_-^T + R_R x_+^T), \tag{36}$$

where $\bar{L}_{L,R}$ are given in (20). Using the above relations, together with $\bar{L}_{L,R}\bar{U} = \hat{g}_{L,R}$ from (3), the boundary terms in (17) are rewritten as

$$\begin{aligned} BT_L &= \left(x_-^T\bar{U} - \mathcal{C}_L^{-1}R_L^*\Lambda_+\tilde{g}_L\right)^* \mathcal{C}_L \left(x_-^T\bar{U} - \mathcal{C}_L^{-1}R_L^*\Lambda_+\tilde{g}_L\right) \\ &\quad + \tilde{g}_L^*\left(\Lambda_+ - \Lambda_+R_L\mathcal{C}_L^{-1}R_L^*\Lambda_+\right)\tilde{g}_L \\ BT_R &= -\left(x_+^T\bar{U} - \mathcal{C}_R^{-1}R_R^*\Lambda_-\tilde{g}_R\right)^* \mathcal{C}_R \left(x_+^T\bar{U} - \mathcal{C}_R^{-1}R_R^*\Lambda_-\tilde{g}_R\right) \\ &\quad - \tilde{g}_R^*(\Lambda_- - \Lambda_-R_R\mathcal{C}_R^{-1}R_R^*\Lambda_-)\tilde{g}_R \end{aligned} \tag{37}$$

where $\mathcal{C}_L = R_L^*\Lambda_+R_L + \Lambda_-$ and $\mathcal{C}_R = \Lambda_+ + R_R^*\Lambda_-R_R$ and $\tilde{g}_L = J_L^{-1}\hat{g}_L$ and $\tilde{g}_R = J_R^{-1}\hat{g}_R$. For an energy estimate of the continuous problem $\mathcal{C}_L \leq 0$ and $\mathcal{C}_R \geq 0$ are necessary in (37).

Proposition VI.5. *The scalar $\mathcal{C}_L$ in* (37) *is non-positive, and hence the non-reflecting boundary condition* (14) *at the left boundary leads to an energy estimate.*

Proposition VI.6. *The matrix $\mathcal{C}_R$ in* (37) *is non-negative, and hence the non-reflecting boundary condition* (14) *at the right boundary leads to an energy estimate.*

The proofs of Proposition VI.5 and Proposition VI.6 are omitted here, but can be found in.[8]

2. Choice of penalty parameters for the discrete formulation

The penalty parameters are

$$\tau_0 = \begin{bmatrix} \tau_0^{11} & \tau_0^{12} \\ \tau_0^{21} & \tau_0^{22} \end{bmatrix}, \qquad \sigma_0 = \begin{bmatrix} \sigma_0^{11} & \sigma_0^{12} \\ \sigma_0^{21} & \sigma_0^{22} \end{bmatrix}, \qquad \tau_N = \begin{bmatrix} \tau_N^{11} \\ \tau_N^{21} \end{bmatrix}, \qquad \sigma_N = \begin{bmatrix} \sigma_N^{11} \\ \sigma_N^{21} \end{bmatrix}. \tag{38}$$

Proposition VI.7. *Choosing the penalty parameter elements τ_0^{ij} and σ_0^{ij} in (38) as*

$$\sigma_0^{11} = \sigma_0^{12} = 0, \qquad \begin{bmatrix} \tau_0^{11} & \tau_0^{12} \\ \tau_0^{21} & \tau_0^{22} \\ \sigma_0^{21} & \sigma_0^{22} \end{bmatrix} = -x_+ \Lambda_+ J_L^{-1}$$

results in a strongly stable numerical scheme.

Proof. Inserting the specific choice above into (30) yields

$$\begin{aligned} BT_L^D = &\left(x_-^T \bar{V}_0 - \mathcal{C}_L^{-1} R_L^* \Lambda_+ \tilde{g}_L\right)^* \mathcal{C}_L \left(x_-^T \bar{V}_0 - \mathcal{C}_L^{-1} R_L^* \Lambda_+ \tilde{g}_L\right) \\ &+ \tilde{g}_L^* \left(\Lambda_+ - \Lambda_+ R_L \mathcal{C}_L^{-1} R_L^* \Lambda_+\right) \tilde{g}_L \\ &- (\bar{L}_L \bar{V}_0 - \hat{g}_L)^* J_L^{-*} \Lambda_+ J_L^{-1} (\bar{L}_L \bar{V}_0 - \hat{g}_L) \end{aligned} \tag{39}$$

where, according to Proposition VI.5, $\mathcal{C}_L \leq 0$. □

Proposition VI.8. *Choosing the penalty parameter elements τ_N^{ij} and σ_N^{ij} in (38) as*

$$\sigma_N^{11} = 0, \qquad \begin{bmatrix} \tau_N^{11} \\ \tau_N^{21} \\ \sigma_N^{21} \end{bmatrix} = x_- \Lambda_- J_R^{-1}$$

results in a strongly stable numerical scheme.

Proof. Inserting the penalty parameters above into (31), yields

$$\begin{aligned} BT_R^D = &- \left(x_+^T \bar{V}_N - \mathcal{C}_R^{-1} R_R^* \Lambda_- \tilde{g}_R\right)^* \mathcal{C}_R \left(x_+^T \bar{V}_N - \mathcal{C}_R^{-1} R_R^* \Lambda_- \tilde{g}_R\right) \\ &- \tilde{g}_R^* \left(\Lambda_- - \Lambda_- R_R \mathcal{C}_R^{-1} R_R^* \Lambda_-\right) \tilde{g}_R \\ &+ (\bar{L}_R \bar{V}_N - \hat{g}_R)^* J_R^{-*} \Lambda_- J_R^{-1} (\bar{L}_R \bar{V}_N - \hat{g}_R) \end{aligned} \tag{40}$$

where $\mathcal{C}_R \geq 0$ according to Proposition VI.6. □

In the propositions above we have used $\bar{V}_0 = [\hat{p}_0, \hat{u}_0, (D\hat{u})_0]^T$ and $\bar{V}_N = [\hat{p}_N, \hat{u}_N, (D\hat{u})_N]^T$.

Remark: Note that when using the penalty parameters as specified in Proposition VI.7 and VI.8, the discrete boundary terms $BT_{L,R}^D$ in (39) and (40) correspond exactly to the continuous boundary terms $BT_{L,R}$ in (37), except for a small damping term. The damping term is a function of the deviation from the boundary data, and goes to zero as the mesh is refined.

VII. Implementation details

Here we describe the numerical procedure, including how the Laplace transform is inverted. As an example, we consider imposing the Dirichlet boundary conditions at the left boundary, and using the exact NRBC at the right boundary. Hence the term $(\Sigma_0 * V)(t) = \mathcal{L}^{-1}\{\hat{\Sigma}_0(s)\hat{V}(s)\}$ in (23) will be replaced by

$$(P^{-1} e_0 \otimes \tau_0^{Dir.} + P^{-1} D^T e_0 \otimes \sigma_0^{Dir.})(L_L V_0 - g_L). \tag{41}$$

Giving Dirichlet boundary conditions such that $p = g_1$ and $u = g_2$ at the left boundary, implies that $L_L = I_2$. The penalty matrices in (41) are chosen such that the numerical scheme becomes stable.

A. Inverting the Laplace transform

At the right boundary we impose the non-reflecting boundary conditions. The convolution $(\Sigma_N * V)(t) = \mathcal{L}^{-1}\{\hat{\Sigma}_N(s)\hat{V}(s)\}$ in (23) is defined as

$$\mathcal{L}^{-1}\{\hat{\Sigma}_N(s)\hat{V}(s)\} = \int_0^t \Sigma_N(\tau)V(t-\tau)d\tau. \tag{42}$$

We follow the work in,[21,22] and approximate the integral (42) at time $t_n = nh$ by the convolution quadrature

$$\sum_{j=0}^{n} \omega_j(h)V(t_{n-j}), \tag{43}$$

where h is the time step, and where $\omega_j(h) \approx h\Sigma_N(t_j)$ for jh away from zero. The coefficients $\omega_j(h)$ in (43) are approximated by

$$\hat{\omega}_j(h) = \rho^{-j}\frac{1}{L}\sum_{l=0}^{L-1}\hat{\Sigma}_N\left(\frac{\delta(\rho e^{i\tau_l})}{h}\right)e^{-ij\tau_l}, \qquad \tau_l = 2\pi l/L. \tag{44}$$

The constants ρ and L and the function δ must be suitably chosen. We use $\rho = 0.975$, $L = T/h$, where T is the end time of the computation, and $\delta(\zeta) = \sum_{i=1}^{3}\frac{1}{i}(1-\zeta)^i$.

Remark: Note that there exist more elaborate versions of this method, see e.g.[23]

B. Time discretization

We let the boundary data $\hat{g}_R$ be zero such that $\hat{\Gamma}_N = 0$ in (25) and $\Gamma_N = 0$ in (23). The semi-discrete scheme (23) is then expressed as

$$V_t = \mathbf{F}(t, V), \tag{45}$$

such that

$$\mathbf{F}(t,V) = \mathbf{A}V + \mathbf{G}(t) + \int_0^t \Sigma_N(\tau)V(t-\tau)d\tau, \tag{46}$$

where, including the Dirichlet boundary condition in (41),

$$\begin{aligned}\mathbf{A} &= -(D\otimes A) + (D^2\otimes B) + (P^{-1}e_0\otimes\tau_0^{Dir.} + P^{-1}D^Te_0\otimes\sigma_0^{Dir.})(e_0^T\otimes L_L)\\ \mathbf{G}(t) &= F - (P^{-1}e_0\otimes\tau_0^{Dir.} + P^{-1}D^Te_0\otimes\sigma_0^{Dir.})g_L(t).\end{aligned}$$

The ordinary differential equation (45) is discretized in time using the trapezoidal rule,

$$V_{n+1} = V_n + \frac{h}{2}\left(\mathbf{F}(t_n, V_n) + \mathbf{F}(t_{n+1}, V_{n+1})\right). \tag{47}$$

We insert (46) into (47), and use the approximation

$$\int_0^t \Sigma_N(\tau)V(t-\tau)d\tau \approx \sum_{j=0}^{n}\hat{\omega}_j(h)V(t_{n-j}).$$

After moving all terms containing V_{n+1} to the left-hand side, we obtain the final scheme

$$\begin{aligned}\left(I - \frac{h}{2}(\mathbf{A}+\hat{\omega}_0)\right)V_{n+1} = \left(I + \frac{h}{2}\mathbf{A}\right)V_n + \frac{h}{2}\sum_{j=0}^{n}\left(\hat{\omega}_j + \hat{\omega}_{j+1}\right)V_{n-j}\\ + \frac{h}{2}\left(\mathbf{G}(t_n) + \mathbf{G}(t_{n+1})\right).\end{aligned} \tag{48}$$

When computing $\hat{\omega}_j$ in (48), using (44), we need $\hat{\Sigma}_N$. We rewrite the parts of $\hat{\Sigma}_N \hat{V}$ in (28) such that we can identify

$$\hat{\Sigma}_N = \bar{P}^{-1}(E_N \otimes \tau_N H_R + D^T E_N \otimes \sigma_N H_R + E_N D \otimes \tau_N G_R + D^T E_N D \otimes \sigma_N G_R), \tag{49}$$

where $E_N = e_N e_N^T$. That is, $\hat{\Sigma}_N$ is a $2(N+1) \times 2(N+1)$ matrix, and consequently so are $\hat{\omega}_j$. Fortunately $\hat{\Sigma}_N$ is sparse since E_N mainly consist of zeroes, and it suffice to compute the lower right corner of $\hat{\omega}_j$.

Remark: The scheme (48) exemplifies the special case when having the Dirichlet boundary conditions at the left boundary and the exact NRBC at the right boundary. Other scenarios, for example when having the exact NRBC's at the left boundary and the Dirichlet boundary condition at the right boundary, are derived in a similar way.

VIII. Numerical results

We let the computational domain be $[x_L, x_R] = [0, 1]$, and as reference solution we use the solution from a five times larger domain. The errors are defined as the difference between the solution and the reference solution, as $\Delta p = p - p^{ref}$ and $\Delta u = u - u^{ref}$. The SBP matrix P is used for computing norms of the errors, as

$$\text{Error}(p) = \|\Delta p\|_P, \qquad\qquad \text{Error}(u) = \|\Delta u\|_P,$$

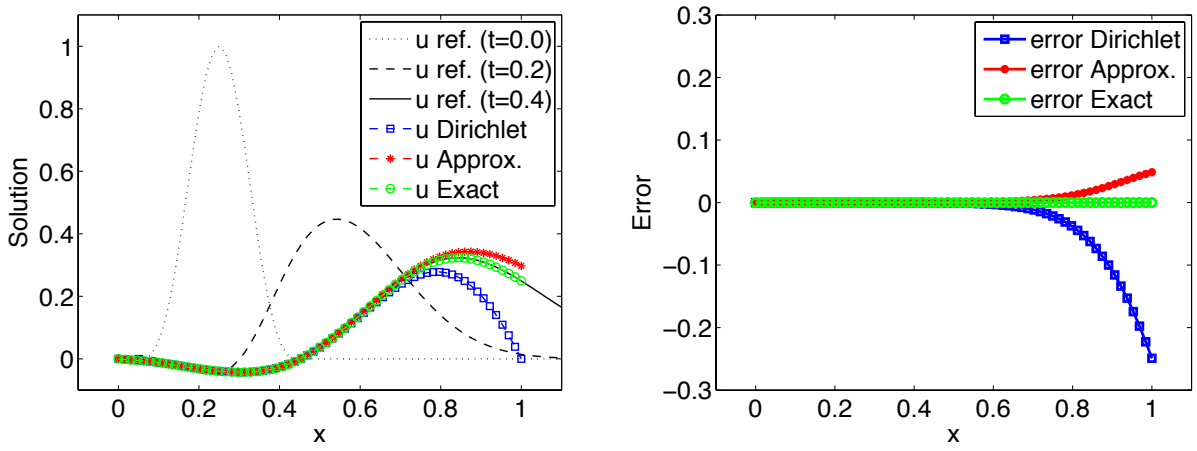
where the norm of a vector $\mathbf{v}$ is defined as $\|\mathbf{v}\|_P^2 = \mathbf{v}^T P \mathbf{v}$. See[20] for details on the accuracy and interpretations of SBP norms. For the space discretization we use a third order accurate SBP scheme, and as mentioned earlier, the trapezoidal rule is used for the time discretization. In all simulations we use the physical parameter values $c = 1$, $v = 0.5$ and $\varepsilon = 0.1$. The time step is $h = 0.001$ and the end time $T = 0.4$. The number of grid point varies, but in the figures we have used $N = 64$. The time step is sufficiently small, such that the errors from the space discretization are dominating. Both the "replacing the indefinte terms" penalty and the "replacing the ingoing variables" penalty have been used in the simulations, and the results are equally good. To reduce the number of figures we only show the solution for the variable u, but the results for the variable p are similar and presented in the tables.

A. Non-reflecting boundary conditions at the right boundary

First, simulations are performed using the scheme (48). As initial condition we use

$$p(x,0) = u(x,0) = \begin{cases} 0 & 0.05 \geq x \\ cos^3(2.5\pi(x - 0.25)) & 0.05 < x < 0.45 \\ 0 & x \geq 0.45. \end{cases} \tag{50}$$

At the left boundary the Dirichlet boundary conditions are imposed and at the right boundary the solution is supposed to propagate out without reflections. This is the same problem setup as in the introducing examples in Figure 1 and Table 1. In comparison the exact NRBC outperforms those examples by far, see Figure 2.



(a) $t = 0.4$ (b) $t = 0.4$

Figure 2. The solution to (1) with initial condition given by (50). At $x = 1$ the pulse should pass without reflections. At the right boundary the Dirichlet boundary condition, the approximate NRBC or the exact NRBC is used, (here with the "replacing the indefinite terms" penalty).

More importantly, the exact NRBC solution converges to the reference solution as the mesh is refined, see Table 2 and 3. The errors have the same size, independently of whether the "replacing the indefinite terms" or the "replacing the ingoing variables" penalty is used. In the simulations, the computational cost when using the exact NRBC's are the same as when using any of the other boundary conditions.

N	Error(p)	ratio	conv. rate	Error(u)	ratio	conv. rate
16	0.00094582			0.00120231		
32	0.00010451	9.0497	3.1779	0.00014386	8.3575	3.0631
64	0.00001193	8.7569	3.1304	0.00001848	7.7862	2.9609
128	0.00000142	8.3796	3.0669	0.00000239	7.7320	2.9508

Table 2. Results obtained using the exact NRBC (with the "replacing the indefinite terms" penalty) at the right boundary.

N	Error(p)	ratio	conv. rate	Error(u)	ratio	conv. rate
16	0.00091109			0.00121302		
32	0.00010158	8.9690	3.1649	0.00014664	8.2722	3.0483
64	0.00001152	8.8217	3.1411	0.00001872	7.8317	2.9693
128	0.00000139	8.2978	3.0527	0.00000241	7.7753	2.9589

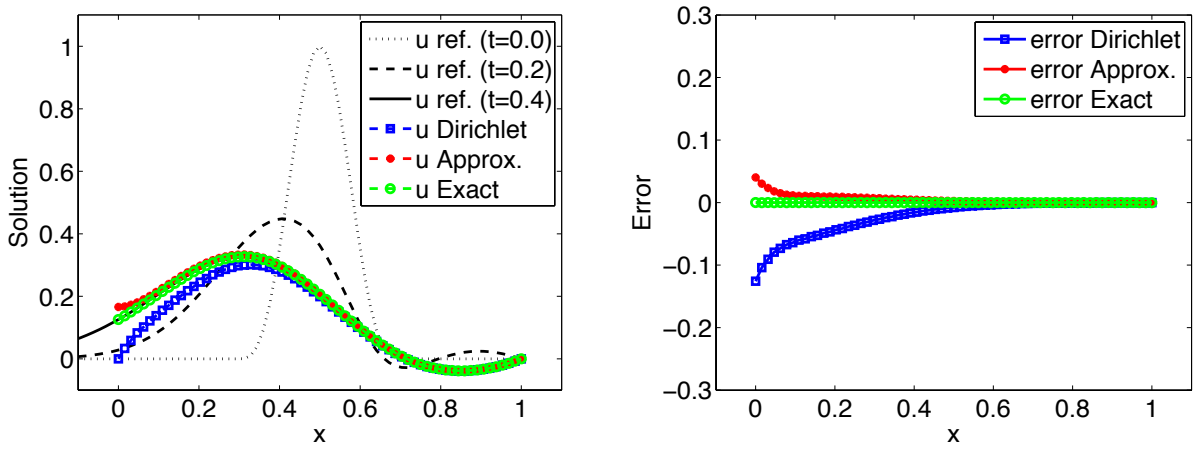
Table 3. Results obtained using the exact NRBC (with the "replacing the ingoing variables" penalty) at the right boundary.

B. Non-reflecting boundary conditions at the left boundary

Next we consider the NRBC's at the left boundary. For this case we use the initial condition

$$p(x,0) = -u(x,0) = \begin{cases} 0 & 0.3 \geq x \\ -cos^3(2.5\pi(x-0.5)) & 0.3 < x < 0.7 \\ 0 & x \geq 0.7, \end{cases} \tag{51}$$

such that the main content of the initial solution travels in the left direction. The resulting solution at time $t = 0.4$ is shown in Figure 3, and as can be seen in Table 4 the solution converges to the reference solution as the mesh is refined.



(a) $t = 0.4$ (b) $t = 0.4$

Figure 3. The solution to (1) with initial condition given by (51). At $x = 0$ the pulse should pass without reflections. At the left boundary the Dirichlet boundary condition, the approximate NRBC or the exact NRBC is imposed (with the "replacing the ingoing variables" penalty).

N	Error(p)	ratio	conv. rate	Error(u)	ratio	conv. rate
16	0.00026816			0.00036867		
32	0.00003824	7.0134	2.8101	0.00005167	7.1355	2.8350
64	0.00000414	9.2323	3.2067	0.00000522	9.9028	3.3078
128	0.00000051	8.0689	3.0124	0.00000064	8.1321	3.0236

Table 4. Results obtained using the exact NRBC (with the "replacing the ingoing variables" penalty) at the left boundary.

The results obtained using the "replacing the indefinite terms" are omitted since those are similar to the results obtained using the "replacing the ingoing terms" penalty.

C. Initial condition without compact support

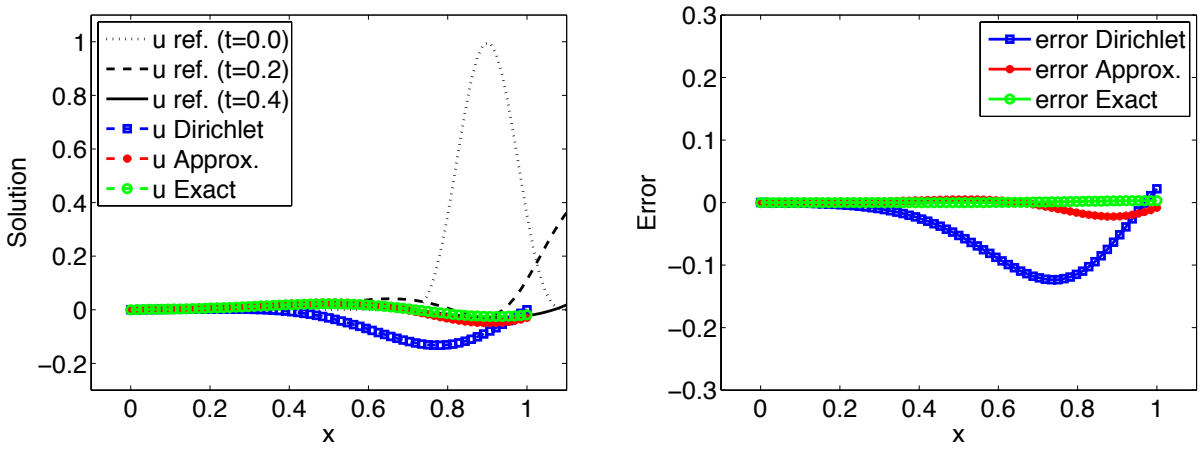
In the boundary conditions (14) the possibility of perturbed data, due to an unknown particular solution, is indicated. To model this, we also use an initial condition that does not have compact support in $x \in [0, 1]$,

$$p(x,0) = u(x,0) = \begin{cases} 0 & 0.7 \geq x \\ cos^3(2.5\pi(x-0.9)) & 0.7 < x < 1.1 \\ 0 & x \geq 1.1, \end{cases} \tag{52}$$

where $p(1,0) = u(1,0) \approx 0.35$. Despite this, we still give zero boundary data to the non-reflecting boundary condition (which we know is wrong, i.e. g'_R in (14) will be non-zero). The results for the exact NRBC's are still superior compared to the ones obtained with the Dirichlet or the approximate NRBC's, see Figure 4. However, since the boundary data does not match the non-zero particular solution, the convergence rates are zero.

IX. Conclusions

We have investigated exact non-reflecting boundary conditions (NRBC) for flow problems, with focus on the theoretical aspects, well-posedness and stability. We consider an incompletely parabolic system of partial differential equations, as a model of the Navier-Stokes equations. The exact NRBC's were derived in Laplace transformed space.



(a) $t = 0.4$ (b) $t = 0.4$

Figure 4. The solution to (1) with initial condition given by (52). At $x = 1$ the pulse should pass out without reflections. At the right boundary the Dirichlet boundary condition, the approximate NRBC or the exact NRBC is imposed, (here with the "replacing the indefinite terms" penalty).

We express the transformed solution as a superposition of ingoing and outgoing waves, and eliminate the ingoing waves at each boundary. Both inflow and outflow NRBC's are derived. It is shown that the exact non-reflecting boundary conditions lead to well-posedness, both in the GKS sense and in the energy sense.

The system is discretized in space using a high order accurate finite difference scheme on Summation-By-Parts form (SBP), and the boundary conditions are imposed weakly using a penalty formulation (SAT). With the continuous energy estimate as a guideline, two different SAT formulations have been derived, both yielding a discrete energy estimate mimicking the continuous one. Hence, by the combined use of the SBP operators and the SAT implementation, stability follows directly from the result of well-posedness for the continuous problem.

The exact non-reflecting boundary conditions are global in time, and must be transformed back for the numerical experiments. This is done by employing convolution quadratures. In the simulations the solutions converge to a reference solution, as accurately as the design order of the numerical scheme. The two different SAT formulations derived perform equally good, producing almost identical results in the numerical simulations.

We have compared the exact NRBC's to the Dirichlet boundary conditions and to approximate NRBC's. The exact NRBC's outperform the other conditions, yielding lower reflections both for exact and erroneous boundary data. Unlike the approximative non-reflecting boundary conditions and the Dirichlet boundary conditions, the exact ones yields convergence to the correct solution when the mesh is refined (and exact boundary data is available).

The superior accuracy, both on the boundary and in the interior (owing to the exact NRBC's and the high order scheme, respectively), in combination with the guaranteed stability, results in a competitive numerical method for computations on unbounded domains.

A. Multiple roots

We show that the polynomial $q(\kappa, s)$ in (6) has no multiple roots κ for $\mathrm{Re}(s) \geq 0$, unless $v = c$. We start by writing $\tilde{q}(\kappa, s) = -q(\kappa, s)/(\varepsilon v)$ as

$$\tilde{q}(\kappa, s) = \kappa^3 - r_2\kappa^2 + r_1\kappa - r_0$$

where the coefficients r_0, r_1 and r_2 are given in (9). The derivative $\tilde{q}'(\kappa, s) = \frac{\partial}{\partial\kappa}\tilde{q}(\kappa, s)$,

$$\tilde{q}'(\kappa, s) = 3\kappa^2 - 2r_2\kappa + r_1,$$

has roots

$$\kappa_{4,5} = \frac{r_2}{3} \pm \sqrt{\left(\frac{r_2}{3}\right)^2 - \frac{r_1}{3}}.$$

If the polynomial $\tilde{q}(\kappa, s)$ has a multiple root κ_j, then that root κ_j will be a solution to the derivative $\tilde{q}'(\kappa, s)$ as well. To check whether $\tilde{q}(\kappa, s)$ and $\tilde{q}'(\kappa, s)$ have any roots in common, we insert $\kappa_{4,5}$ into $\tilde{q}(\kappa, s)$. This yields

$$\tilde{q}(\kappa_{4,5}, s) = \frac{-1}{27}\left(r_2\left(2r_2^2 - 9r_1\right) \pm 2\sqrt{r_2^2 - 3r_1}\left(r_2^2 - 3r_1\right) + 27r_0\right).$$

Requiring $\tilde{q}(\kappa_{4,5}, s) = 0$ leads to

$$r_2\left(2r_2^2 - 9r_1\right) + 27r_0 = \mp 2\sqrt{r_2^2 - 3r_1}\left(r_2^2 - 3r_1\right),$$

which we square on both sides to obtain

$$\left(r_2\left(2r_2^2 - 9r_1\right) + 27r_0\right)^2 = 4\left(r_2^2 - 3r_1\right)^3. \tag{53}$$

If the relation (53) is fulfilled $q(\kappa, s)$ has a multiple root. We check if this can occur by defining $\Upsilon = \left(r_2\left(2r_2^2 - 9r_1\right) + 27r_0\right)^2 - 4\left(r_2^2 - 3r_1\right)^3$, and see whether it is possible to find $\Upsilon = 0$. Inserting the values $r_0 = s^2/(\varepsilon v)$, $r_1 = -2s/\varepsilon$ and $r_2 = (v^2 - c^2 - s\varepsilon)/(\varepsilon v)$ from (9) gives

$$\Upsilon = -27\frac{s^2}{\varepsilon^4 v^4}\left(4c^2(v^2 - c^2)^2 + 4c^2(3c^2 + 5v^2)s\varepsilon + (v^2 + 12c^2)(s\varepsilon)^2 + 4(s\varepsilon)^3\right).$$

Let $s\varepsilon = \tilde{\eta} + \tilde{\xi}i$ to split Υ into one real and one imaginary part, as

$$\begin{aligned}\Upsilon = &-27\frac{s^2}{\varepsilon^4 v^4}\left(4c^2(v^2 - c^2)^2 + 4c^2(3c^2 + 5v^2)\tilde{\eta} + (v^2 + 12c^2)(\tilde{\eta}^2 - \tilde{\xi}^2) + 4(\tilde{\eta}^3 - 3\tilde{\eta}\tilde{\xi}^2)\right)\\ &-27\frac{s^2}{\varepsilon^4 v^4}\left(4c^2(3c^2 + 5v^2) + 2(v^2 + 12c^2)\tilde{\eta} + 4(3\tilde{\eta}^2 - \tilde{\xi}^2)\right)\tilde{\xi}i.\end{aligned}$$

The imaginary part of Υ can be cancelled either by choosing $\tilde{\xi} = 0$ or by choosing $\tilde{\xi}^2 = c^2(3c^2 + 5v^2) + (v^2 + 12c^2)\tilde{\eta}/2 + 3\tilde{\eta}^2$. In both these cases the real part of Υ can only be cancelled if $\tilde{\eta} < 0$. The only exception is if $s = 0$, then a multiple root is possible. That case must be treated separately, and the matrix $E(s)$ in Proposition IV.1 is in fact non-singular unless $s = 0$ and $v = c$. In this paper we will simply avoid the special case $v \neq c$.

References

[1]D. Appelö, T. Hagstrom, and G. Kreiss. Perfectly matched layers for hyperbolic systems: general formulation, well-posedness, and stability. *SIAM J. Appl. Math.*, 67(1):1–23, 2006.

[2]E. Bécache, D. Givoli, and T. Hagstrom. High-order absorbing boundary conditions for anisotropic and convective wave equations. *Journal of Computational Physics*, 229(4):1099–1129, 2010.

[3]J.P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2):185–200, 1994.

[4]M. H. Carpenter, D. Gottlieb, and S. Abarbanel. Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes. *Journal of Computational Physics*, 111(2):220–236, 1994.

[5]M.H. Carpenter, J. Nordström, and D. Gottlieb. A stable and conservative interface treatment of arbitrary spatial accuracy. *Journal of Computational Physics*, 148:341–365, 1999.

[6]B. Engquist and B. Gustafsson. Steady state computations for wave propagation problems. *Mathematics of Computations*, 49:39–64, 1987.

[7]B. Engquist and A. Majda. Absorbing boundary conditions for the numerical simulation of waves. *Mathematics of Computation*, 31(139):629–651, 1977.

[8]Sofia Eriksson. *Stable Numerical Methods with Boundary and Interface Treatment for Applications in Aerodynamics*. PhD thesis, Uppsala University, Division of Scientific Computing, Numerical Analysis, No. 985, ISBN 978-91-554-8509-2, 2012.

[9]L. Ferm. Non-reflecting boundary conditions for the steady Euler equations. *Journal of Computational Physics*, 122:307–316, 1995.

[10]M. Grote and J. Keller. Exact nonreflecting boundary conditions for the time dependent wave equation. *SIAM Journal on Applied Mathematics*, 55(2):280–297, 1995.

[11]M. J. Grote and J. B. Keller. Nonreflecting boundary conditions for time-dependent scattering. *Journal of Computational Physics*, 127(1):52–65, 1996.

[12]B. Gustafsson. Far-field boundary conditions for time-dependent hyperbolic systems. *SIAM J. ScI. Statist. Comput.*, 9(5):812–828, 1988.

[13]B. Gustafsson and H.-O. Kreiss. Boundary conditions for time dependent problems with an artificial boundary. *Journal of Computational Physics*, 30(3):333–351, 1979.

[14]B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time Dependent Problems and Difference Methods*. John Wiley & Sons, Inc., 1995.

[15]B. Gustafsson, H.-O. Kreiss, and A. Sundström. Stability theory of difference approximations for mixed initial boundary value problems. II. *Mathematics of Computation*, 26(119):649–686, 1972.

[16]T. Hagstrom. Radiation boundary conditions for the numerical simulation of waves. *Acta Numerica*, 8:47–106, 1999.

[17]T. Hagstrom, E. Bécache, D. Givoli, and K. Stein. Complete radiation boundary conditions for convective waves. *Commun. Comput. Phys.*, 11(2):610–628, 2012.

[18]L. Halpern. Artificial boundary conditions for incompletely parabolic perturbations of hyperbolic systems. *SIAM Journal on Mathematical Analysis*, 22(5):1256–1283, 1991.

[19]J.S. Hesthaven. On the analysis and construction of perfectly matched layers for the linearized euler equations. *Journal of Computational Physics*, 142(1):129–147, 1998.

[20]J.E. Hicken and D.W. Zingg. Summation-by-parts operators and high-order quadrature. *Journal of Computational and Applied Mathematics*, 237(1):111–125, 2013.

[21]C. Lubich. Convolution quadrature and discretized operational calculus. I. *Numerische Mathematik*, 52:129–145, 1988.

[22]C. Lubich. Convolution quadrature and discretized operational calculus. II. *Numerische Mathematik*, 52:413–425, 1988.

[23]C. Lubich and A. Schädle. Fast convolution for nonreflecting boundary conditions. *SIAM J. Sci. Comput.*, 24(1):161–182, 2002.

[24]K. Mattsson. Boundary procedures for summation-by-parts operators. *Journal of Scientific Computing*, 18(1):133–153, 2003.

[25]J. Nordström. The influence of open boundary conditions on the convergence to steady state for the Navier-Stokes equations. *Journal of Computational Physics*, 85:210–244, 1989.

[26]J. Nordström and M. H. Carpenter. Boundary and interface conditions for high order finite difference methods applied to the Euler and Navier-Stokes equations. *Journal of Computational Physics*, 148:621–645, 1999.

[27]J. Nordström, S. Eriksson, and P. Eliasson. Weak and strong wall boundary procedures and convergence to steady-state of the Navier-Stokes equations. *Journal of Computational Physics*, 231(14):4867–4884, 2012.

[28]J. Nordström, J. Gong, E. van der Weide, and M. Svärd. A stable and conservative high order multi-block method for the compressible Navier-Stokes equations. *Journal of Computational Physics*, 228(24):9020–9035, 2009.

[29]L. Råde and B. Westergren. *Mathematics Handbook for Science and Engineering*. Studentlitteratur, Lund, 1998.

[30]I. L. Sofronov. Non-reflecting inflow and outflow in a wind tunnel for transonic time-accurate simulation. *Journal of Mathematical Analysis and Applications*, 221(1):92–115, 1998.

[31]J. C. Strikwerda. Initial boundary value problems for incompletely parabolic systems. *Communications on Pure and Applied Mathematics*, 30(6):797–822, 1977.

[32]M. Svärd, M.H. Carpenter, and J. Nordström. A stable high-order finite difference scheme for the compressible Navier-Stokes equations: far-field boundary conditions. *Journal of Computational Physics*, 225(1):1020–1038, 2007.

[33]S. V. Tsynkov. Numerical solution of problems on unbounded domains. A review. *Appl. Numer. Math.*, 27:465–532, August 1998.