Robust and efficient discretizations of wave-dominated problems

GUSTAV ERIKSSON
Abstract

Partial differential equations appear in mathematical models that describe a wide range of physical phenomena, such as sound pressure waves in the air, the vibrations of solid structures, and the flow of fluids. Unfortunately, most of these problems can not be solved analytically using pen and paper. Instead, we turn to numerical methods and computer simulations to obtain approximate solutions. In this thesis, the focus is on high-order accurate finite difference methods for wave propagation and fluid dynamical problems. High-order finite difference methods are conceptually simple to design and implement efficiently on modern computers. However, special care must be taken close to boundaries to obtain robust and stable schemes. In this thesis, a class of finite difference operators with summation-by-parts (SBP) properties is used. These operators satisfy a discrete equivalent to intergration-by-parts which, when the boundary conditions are correctly imposed, enables a stability proof for the discretized scheme. Two such methods for imposing boundary conditions are studied and compared in the thesis, the simultaneous-approximation-term (SAT) method and the projection (P) method.

In Paper I a high-order accurate finite difference discretization of the incompressible Navier-Stokes equations is presented, where the projection method is found to be more suitable for wall boundary conditions. In Paper II the SBP-SAT and SBP-P methods are compared for boundary and interface conditions to the dynamic beam equation and the dynamic Kirchoff-Love plate equation. A new SBP-P and hybrid SBP-P-SAT method is developed for non-conforming interface conditions to the second-order wave equation in Paper III. In Paper IV shape optimization problems constrained by the second-order wave equation are solved using high-order SBP-P-SAT finite difference discretizations. Theoretical aspects of the projection method are discussed in Paper V. In Paper VI SBP operators defined on Gauss-Lobatto quadrature points are used to derive an efficient and robust scheme for the Laplacian on complex geometries.

Keywords: partial differential equations, wave propagation problems, finite difference methods, summation-by-parts, boundary conditions

Gustav Eriksson, Department of Information Technology, Division of Scientific Computing, Box 337, Uppsala University, SE-751 05 Uppsala, Sweden.

© Gustav Eriksson 2024

ISSN 1651-6214
URN urn:nbn:se:uu:diva-527431 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-527431)
List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the publishers.
Contents

1 Introduction ........................................................................................................... 7
2 Summation-by-parts finite differences .............................................................. 12
  2.1 First derivative SBP operators ...................................................................... 12
  2.2 Second derivative SBP operators ................................................................. 15
  2.3 Spatial discretization .................................................................................. 16
3 Temporal discretization ...................................................................................... 19
4 Accuracy and convergence ............................................................................... 20
5 Summary of papers .......................................................................................... 21
6 Sammanfattning på svenska ............................................................................ 25
References ........................................................................................................... 28
1. Introduction

There is a large number of physical processes in the world that can be modeled mathematically by partial differential equations (PDEs). A significant class of such PDEs are wave propagation problems, which can for example be used to model acoustic or electromagnetic waves in the atmosphere and under water, elastic waves in the soil, and vibrations in solid beams and plates. Another important class of PDEs are those used to model fluid dynamical problems. With the advent of computers in the middle of the last century, computational fluid dynamics (CFD) as a scientific field grew rapidly and is used extensively today in the vehicle industry, for example. Being able to solve these types of PDEs accurately and reliably is of crucial importance in science and industry.

As an example model consider a perfectly flexible string of length $L$ that vibrates freely after being disturbed from its equilibrium with the left end fixed and the right end loosely hanging. If $u(x,t)$ denotes the deviation from the equilibrium at position $x$ and time $t$, the motion of the string can be described by the following equations:

\[
\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} , \quad x = [0, L], \quad t > 0, \\
u(0) = 0, \quad x = 0, \quad t > 0, \\
\frac{\partial u}{\partial x} (L, t) = 0, \quad x = L, \quad t > 0, \\
u(0, x) = f(x), \quad x = [0, L], \quad t = 0, \\
u_t (0, x) = 0, \quad x = [0, L], \quad t = 0, 
\]

where $c$ is a constant parameter (wave speed), $f(x)$ is the shape of the string when it is released (initial data), and $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial t}$ denote partial derivatives with respect to space and time, respectively. The first equation in (1.1) is the PDE. The second and third equations are referred to as Dirichlet and Neumann boundary conditions, respectively. The fourth and fifth equations are the initial conditions. Equations of this kind, i.e. a PDE along with both boundary and initial conditions, are often referred to as initial-boundary-value problems (IBVPs). An important concept regarding the solvability of IBVPs is well-posedness. Coined by Hadamard, an IBVP is said to be well-posed if the following holds:

- A solution exists.
- The solution is unique.
• The solution depends smoothly on the data in the problem (the wave speed $c$ and initial data $f(x)$ in (1.1)).

The first two conditions are reasonable for any problem we wish to obtain a solution to and the third condition guarantees that the influence of errors in the data or errors due to finite precision computations is limited. A useful mathematical tool for proving that an IBVP satisfies the third condition is the energy method [7]. The energy method is used repeatedly throughout the thesis and is therefore briefly described below for the example problem (1.1).

Multiplying the first equation in (1.1) by $u_t$, where the subscript denotes partial differentiation, and integrating in space on the interval $[0, L]$ results in

$$
\int_0^L u_t u_{tt} \, dx = c^2 \int_0^L u_t u_{xx} \, dx. \tag{1.2}
$$

Recall the integration-by-parts formula, which states that two continuously differentiable functions $f(x)$ and $g(x)$ defined on the interval $[a, b]$ satisfies

$$
\int_a^b f(x)g'(x) \, dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x) \, dx. \tag{1.3}
$$

By integrating the right-hand-side (RHS) of (1.2) in parts and rearranging, we get

$$
\int_0^L u_t u_{tt} + c^2 u_x u_x \, dx = c^2 u_t u_x|_{x=L} - c^2 u_t u_x|_{x=0} = 0, \tag{1.4}
$$

where the boundary conditions have been inserted in the last step. To avoid explicitly writing out the integrals, the following notation for the $L_2$-inner product and norm of two functions $u, v$ is usually introduced when analyzing IBVPs:

$$
(u, v)_\Omega = \int_\Omega uv \, dx \quad \text{and} \quad ||u||^2_\Omega = (u, u), \tag{1.5}
$$

where $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, denotes the spatial domain (in (1.1) $\Omega = [0, L]$ and $d = 1$). With this notation, we can write (1.4) as

$$
\frac{dE}{dt} = 0, \tag{1.6}
$$

where $E(t) = ||u_t||^2_\Omega + c^2 ||u_x||^2_\Omega$, is a mathematical energy (hence the name of the method) that is conserved over time. Integrating from the initial time to some final time $t > 0$ gives

$$
E(t) = E(0) = c^2 ||f_x||^2_\Omega, \tag{1.7}
$$

i.e. the total mathematical energy in the system remains the same for all time. The energy equation (1.7) is enough to show that the solution $u(x, t)$ to (1.1) depends smoothly on the data [7].
If we can solve (1.1) for the unknown function $u(x,t)$ we know the precise behavior of the string for all time. However, in many situations of practical relevance, all the parameters and functions of IBVPs necessary to obtain a unique solution are not known, but we might have access to measurement data that provides the solution at certain positions in time and space. Consider for example the case where the wave speed $c$ in (1.1) is unknown, but the position of the string at certain locations and times has been recorded. A question then arises as to whether or not the wave speed can be found by just using prior knowledge of the physics (the IBVP) and the measurements. These types of problems are referred to as inverse problems and are usually much harder to solve than IBVPs, but their usefulness in practice is significant. Successfully solving inverse problems can for example help to answer questions such as "How should an aircraft wing be designed to maximize lift?" or "Where should the wind turbines be placed to minimize the noise pollution?". In this thesis, a shape optimization problem is solved in Paper IV, which is a special type of inverse problem where the spatial domain is the unknown quantity.

Although the relatively simple IBVP (1.1) can be solved analytically with pen and paper, modifying it slightly so that, for example, the wave speed is a spatially dependent function complicates this analysis significantly. Most IBVPs used to model physical systems are too complicated to allow for straightforward derivation of closed-form analytical solutions. Instead, we rely on numerical methods and computers to obtain approximate solutions. Many different numerical methods exist to choose from, all with different pros and cons. A practically useful numerical method should be accurate, robust, flexible, and suitable for implementation on modern hardware.

The accuracy of a numerical method is usually evaluated in relation to the computational costs of applying it. A method is of little practical use if only highly accurate solutions can be obtained at unreasonably large computational costs. Instead, we want methods that can provide sufficiently accurate solutions at low costs. Most traditional methods, such as the finite element method (FEM), finite differences (FD), finite volume (FV), or discontinuous Galerkin methods (DG) all split the computational domain into smaller constituents and approximate the continuous operations, such as differentiation and integration, locally. This is often referred to as discretization. For these methods, a more accurate solution can be obtained by simply decomposing the domain into smaller pieces (mesh refinement) or by using a more sophisticated approximation of the continuous operations (higher order of accuracy), both with an increase in computational costs. There exists numerical methods that are conceptually different from the traditional methods, but the pattern remains the same; higher accuracy comes at the price of increased computational costs. Consider for example a very recent approach using physics-informed neural networks (PINNs) [27], where the accuracy of the model is in large part determined by the training process of the neural networks. Longer training on
more data typically leads to a more accurate model but naturally requires more computational resources.

The robustness and flexibility of a numerical method are crucial for anyone using the method to perform numerical experiments. Consider for example an engineer using computer simulations (numerical methods) as a substitute for physical experiments when designing a new vehicle. In these cases, it is not only important that the numerical method provides accurate representations of the physical world, but the method must also work well on all variations of the vehicle and surrounding physics the engineer is interested in. For a numerical method solving the IBVP in (1.1), an engineer might, for example, be interested in adding a term to the RHS of the PDE corresponding to gravity or friction or in changing one of the boundary conditions to a mixed-type boundary condition (where the end of the string is attached to a spring). Whatever numerical method is used to solve (1.1), it should ideally be able to handle these modifications without much trouble. An important concept that is related to the robustness of a numerical method is stability. A provably stable numerical method will almost per definition be robust, as a mathematical proof exists that guarantees that the approximate solution can not grow unphysically. A useful property of many numerical methods that can be used when analyzing stability is summation-by-parts (SBP). Named so for its similarities to integration-by-parts, the SBP property can be used to perform a discrete analog to the analysis done using the energy method in (1.2)-(1.7). As we shall see in Section 2.3, how to impose the boundary conditions in the discrete context is of crucial importance to fully utilize the SBP properties. Many traditional numerical methods that are widely used today can be written in a form such that they satisfy the SBP property, such as FEM, FV, and DG [23, 2, 5, 4].

As previously mentioned, it is central for a numerical method’s success that the computational costs of using the method are low, which naturally is closely related to the computer implementation of the method. Historically, a lot of interest was given to the number of floating-point operations (FLOPS) required by a numerical method, as most early computers were single-threaded (meaning that operations happen sequentially). Today this measure is still relevant, but perhaps more important is how parallelizable a numerical method is. Most devices in use by average people today have multiple computational cores and in the top echelon of computational power, we have supercomputers that can consist of millions of cores coupled together. To fully utilize these devices, it is therefore important that a numerical method can be parallelized efficiently. It is also important to consider the complexity of the implementation. The work needed to develop and maintain the computer code should be kept as little as possible, which is easier for more straightforward and intuitive implementations. Furthermore, simple parallel algorithms often lend themselves to efficient implementations on graphics processing units (GPUs), which are very attractive from a performance-per-watt point of view.
In this thesis, the main focus is on finite difference methods. These methods are conceptually different from Galerkin methods (FEM, DG, etc.) as they directly discretize the IBVP (1.1) (the strong form) rather than the integrated equation (the weak form). The main benefits of using finite differences are that high-order discretizations can be obtained relatively easily and that they are straightforward to implement on parallel computational platforms with good performance. Additionally, most of the IBVPs considered in the thesis are wave propagation problems (Papers II-VI), for which high-order methods are known to be very efficient [10]. In the past, finite differences were limited to low orders of accuracy to guarantee robustness, but with the advent of SBP finite difference operators, this is no longer a major issue. However, special care must be taken to not lose stability when imposing boundary conditions. How to do this for wave propagation problems using SBP finite differences is one of the main themes throughout the thesis. In addition to that, the incompressible Navier-Stokes equations are solved in Paper I (CFD), which also benefits from high-order methods to resolve intricate turbulence-like effects. And, in paper VI, a spectral element SBP method is considered for the second-order wave equation. This method is conceptually different from finite difference methods, but since it is formulated precisely as a SBP finite difference method, it will not be treated in more detail in this summary.

The majority of this thesis consists of six papers (three published and three submitted manuscripts). An overview of basic concepts used in the papers is presented in the upcoming chapters (Chapters 2-4). More details and discussions on novelties can be found in the individual papers. In Chapter 2 the SBP finite difference method is introduced. As an illustrating example, the spatial discretization of the IBVP (1.1) is considered. In Chapter 3 the temporal discretizations used in the papers are briefly discussed and in Chapter 4 some remarks on the accuracy properties of these discretizations are included. Short summaries of each paper are presented in Chapter 5.
2. Summation-by-parts finite differences

The concepts of SBP finite difference methods were first introduced in the 1970s as a remedy to the stability problems of traditional finite difference methods for IBVPs [11]. In the following decades, sporadic work was done developing the theoretical framework of the methods, but it was not until the 1990s that they grew in popularity substantially with the introduction of two different stable ways of imposing boundary conditions (discussed in Section 2.3). Since then SBP finite differences have been used to solve a wide range of IBVPs and the SBP concepts have even been extended to methods outside the finite difference community. In this chapter, some basic theory on standard one-dimensional SBP finite differences is presented. More complicated applications of these concepts that are used in the thesis, such as multidimensional discretizations on non-rectangular domains and inverse problems, are described in the individual papers.

2.1 First derivative SBP operators

Consider a decomposition of the interval \( [0, L] \) into \( m \) equidistant grid points \( x_1 < x_2 < \ldots < x_m \). In this thesis, only operators that include the endpoints are considered, i.e. \( x_1 = 0 \) and \( x_m = L \), but there exist generalized SBP operators where this is not a limitation [4]. There also exists SBP operators that are defined on grids with non-equidistant points, such as the so-called boundary-optimized SBP operators or operators defined on spectral collocation points (which are used in Paper I and VI, respectively). For a function \( f(x) \) on the interval \([0, L] \), let \( f = [f_1, f_2, \ldots, f_m]^\top \in \mathbb{R}^m \) denote a column vector restricting the function to the grid such that \( f_i = f(x_i), \ i = 1, 2, \ldots, m \). As an example illustrating the SBP concept, consider the well-known second-order accurate central finite difference approximation of the first derivative of \( f(x_i) \), given by

\[
\frac{df(x_i)}{dx} \approx \frac{f(x_{i+1}) - f(x_{i-1})}{2h}, \quad i = 2, 3, \ldots, m-1, \quad (2.1)
\]

where \( h = \frac{L}{m-1} \) is the step size. The second-order accuracy means that the error of the approximation in (2.1) is proportional to \( h^2 \) for smooth functions, i.e. by halving the step size \( h \) the error decreases by one-fourth. On the boundaries \( x = x_1 \) and \( x = x_m \) the second-order approximation does not fit. Instead, the
following first-order accurate forward and backward difference formulas can be used:

\[
    f'(x_1) \approx \frac{f(x_2) - f(x_1)}{h} \quad \text{and} \quad f'(x_m) \approx \frac{f(x_m) - f(x_{m-1})}{h}.
\]  

(2.2)

The finite difference stencils in (2.1) and (2.2) for all grid points can be concisely written as \( \mathbf{D}_1 \mathbf{f} \approx f'(x) \), where \( \mathbf{D}_1 \) is a first derivative approximation matrix given by

\[
\mathbf{D}_1 = \frac{1}{h} \begin{bmatrix}
    -1 & 1 \\
    -\frac{1}{2} & 0 & \frac{1}{2} \\
    \ddots & \ddots & \ddots \\
    -\frac{1}{2} & 0 & \frac{1}{2} \\
    -1 & 1
\end{bmatrix}.
\]  

(2.3)

The derivative approximation matrix \( \mathbf{D}_1 \) is the first component needed to obtain a discrete equivalent to integration-by-parts, the second is a discrete equivalent to the \( L_2 \)-inner product. Let \( \mathbf{H} = \mathbf{H}^\top > 0 \) denote a \( m \times m \) symmetric positive definite (SPD) matrix and define a discrete inner product as

\[
    (\mathbf{u}, \mathbf{v})_\mathbf{H} = \mathbf{u}^\top \mathbf{H} \mathbf{v}, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^m,
\]  

(2.4)

and corresponding norm \( \|\mathbf{u}\|_\mathbf{H}^2 = (\mathbf{u}, \mathbf{u})_\mathbf{H} \). In the finite difference community, the matrix \( \mathbf{H} \) is often referred to as the norm matrix or simply the norm. The condition on \( \mathbf{D}_1 \) and \( \mathbf{H} \) to satisfy the SBP property is given in the following definition:

**Definition 2.1.1.** A difference operator \( \mathbf{D}_1 \approx \frac{\partial}{\partial x} \) is said to be a first derivative SBP operator if, for the norm matrix \( \mathbf{H} \), the relation

\[
    \mathbf{H} \mathbf{D}_1 + \mathbf{D}_1^\top \mathbf{H} = -\mathbf{e}_1 \mathbf{e}_1^\top + \mathbf{e}_m \mathbf{e}_m^\top,
\]  

(2.5)

holds, where \( \mathbf{e}_1, \mathbf{e}_m \in \mathbb{R}^m \) are boundary restriction operators given by

\[
\mathbf{e}_1 = \begin{bmatrix}
    1 \\
    0 \\
    \vdots \\
    0
\end{bmatrix} \quad \text{and} \quad \mathbf{e}_m = \begin{bmatrix}
    0 \\
    0 \\
    \vdots \\
    1
\end{bmatrix}.
\]  

(2.6)

The SBP property (2.5) is equivalent to

\[
    (\mathbf{u}, \mathbf{D}_1 \mathbf{v})_\mathbf{H} = \mathbf{u}_m \mathbf{v}_m - \mathbf{u}_1 \mathbf{v}_1 - (\mathbf{D}_1 \mathbf{u}, \mathbf{v})_\mathbf{H}, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^m,
\]  

(2.7)

since e.g. \( \mathbf{e}_1^\top \mathbf{u} = u_1 \). Note the similarity between (2.7) and its continuous counterpart in (1.3) (integration-by-parts).
As it turns out, the structure of $H$ has significance for the stability proofs and the accuracy of SBP finite difference discretizations \cite{16, 11, 12}. In this thesis, SBP operators with diagonal $H$ are exclusively considered since these allow for stable discretizations of non-linear, variable coefficient, and coordinate transformed IBVPs \cite{30, 14, 2}. With $D_1$ given by (2.3), the SBP property is satisfied for the diagonal inner product matrix

\[
H = h \begin{bmatrix}
\frac{1}{2} & 1 & & \\
& \ddots & \ddots & \\
& & 1 & \frac{1}{2}
\end{bmatrix},
\]  

(2.8)

Note that $(u, v)_H$ with $H$ given by (2.8) is precisely the trapezoidal rule for the integral $\int_0^L u(x)v(x) \, dx$ (the $L_2$-inner product). This means that the quadrature rule defined by $H$ and the difference stencils in the interior of $D_1$ are second-order accurate, while the boundary stencils of $D_1$ are only first-order accurate. This is typical for SBP finite difference operators. For diagonal inner product matrices in particular, it can be shown that the accuracy of the boundary stencils can be at most half the accuracy of the interior stencils for the SBP property to hold \cite{12}. This reduction in accuracy for diagonal inner product matrices will affect the overall convergence rate of the discretized IBVP, which is discussed in more detail in Chapter 4.

The matrices $D_1$ and $H$ in (2.3) and (2.8) are perhaps the simplest finite difference SBP operators one can imagine and were presented already in the pioneering paper in 1974 \cite{11}. In general, finding practically useful $D_1$ and $H$ matrices that satisfy (2.5) is non-trivial, especially if high orders of accuracy are desired. Typically, the interior stencils of $D_1$ can be determined using Taylor expansions, but finding the boundary stencils and entries in $H$ is tedious and usually involves some amount of numerical optimization. Important aspects to consider are the following:

- The accuracy of $D_1$ and $H$, usually measured in terms of the highest degree polynomial they can differentiate and integrate exactly.
- The spectrum of $D_1$, i.e. the distribution of eigenvalues of the discretization matrix. A too large spectral radius (largest eigenvalue in magnitude) will negatively impact the efficiency of time stepping schemes and iterative solvers.
- The number of non-zeros in the stencils, to keep the number of floating-point operations when applying the operators as low as possible.

Nevertheless, there exists today a slew of SBP finite difference operators with different structures and accuracies, adapted for different IBVPs. For example, the upwind SBP operators, with skewed stencils and built-in artificial dissipation, have been used to construct robust and efficient schemes for the com-
pressible Euler equations [15]. Another example is the boundary-optimized SBP operators, with non-equidistant grid point distributions close to the boundaries, which have shown to be highly efficient for wave propagation problems and problems where effects close to boundaries are of primary interest (see e.g. [18, 17, 29] or Paper I).

2.2 Second derivative SBP operators

As many wave propagation and CFD problems include second derivative terms, it is imperative that we also have second derivative SBP operators to approximate them, based on the same norm $H$. An immediate candidate to approximate second derivatives is to use the first derivative operator twice, i.e. $\hat{D}_2 = D_1 D_1$. However, this operator has some drawbacks. First of all, the finite difference stencils in $\hat{D}_2$ are unnecessarily wide, which increases the computational costs of applying them. This can be realized by considering the second-order central finite difference approximation in (2.1) (by applying it once again to approximate $f''(x_i)$). Furthermore, the accuracy of the boundary stencils in $\hat{D}_2$ is decreased by one compared to $D_1$, and even though the operator is dissipative, it can not be used to dampen spurious oscillations that may occur (the checkerboard pattern sometimes referred to as the $\pi$-mode of the grid) [20]. This can lead to unphysical solutions which require artificial dissipation to rectify. As an alternative to the wide operators, narrow stencil second derivative SBP operators were derived in [22] that retain the boundary accuracy and dampen the spurious oscillations. These operators are defined as follows:

Definition 2.2.1. Let $D_1$ be a first derivative SBP operator. A difference operator $D_2 \approx \frac{\partial^2}{\partial x^2}$ is said to be a compatible second derivative SBP operator if, for the norm $H$, the relation

$$D_2 = H^{-1}(-D_1^T H D_1 - R - e_1 d_1^T + e_m d_m^T),$$

(2.9)

holds, where $d_1$ and $d_m$ are one-sided approximation of the first derivatives at the boundaries and $R = R^T \geq 0$.

The term compatible in Definition 2.2.1 refers to the fact that $D_2$ can be written as $D_2 = \hat{D}_2 - H^{-1}R$, i.e. as the wide operator with a correction term. The term $R$ consists of high-order derivative operators that are chosen to narrow the stencils of $\hat{D}_2$, increase the boundary accuracy, and add dissipation to counteract spurious oscillations. For many IBPVs of practical interest variable-coefficient second derivative terms of the form $\frac{\partial}{\partial \xi}(a(x) \frac{\partial}{\partial \xi})$ are present. One important example is the second-order wave equation in curvilinear coordinates, which is treated in Papers IV and VI. These terms can be
discretized using variable-coefficient second derivative SBP operators, derived for equidistant grids for the first time in [14], with similar SBP properties as in Definition 2.2.1.

2.3 Spatial discretization

We now return to the IBVP (1.1) and the spatial discretization of the problem. Using the SBP operators introduced in the previous section, a consistent spatial discretization of (1.1) is given by

\[
\frac{d^2v}{dt^2} = c^2 D_2 v, \quad t > 0,
\]
\[
e_1^T v = 0, \quad t > 0,
\]
\[
de_m^T v = 0, \quad t > 0,
\]
\[
v = f, \quad t = 0,
\]
\[
\frac{dv}{dt} = 0, \quad t = 0.
\]

(2.10)

Here time is left continuous, so that \( v \) is a vector where each element depends on \( t \). The spatial discretization (2.10) is essentially a second-order ordinary differential equation (ODE) constrained by the discretized boundary conditions (the second and third equations). Instead of trying to solve this constrained system numerically the boundary conditions are usually incorporated into the ODE, resulting in an unconstrained system that can be solved using any suitable ODE solver. How to do this in a way that utilizes the SBP properties of the discretization operators to obtain a provably stable scheme has been a topic of research for a long time. Common finite difference approaches, such as the injection method for example, hinder a proof like that for many types of boundary conditions [13, 31]. In this thesis, two methods for stably imposing the boundary conditions are considered, the simultaneous approximation term method (SBP-SAT) and the projection method (SBP-P).

With the SBP-SAT method a penalty term is added to the ODE that weakly imposes the boundary conditions and, by carefully choosing the parameters in the penalty term, stability can be obtained. The SBP-SAT method was first presented in [3] and has since then been the most studied and developed method for imposing boundary conditions using SBP finite differences. However, the SBP-SAT method can be problematic in certain situations. Often the method introduces free parameters that must be chosen and, for problems involving second derivatives (or higher) and certain boundary conditions, finding parameters that lead to stable schemes can be difficult [22, 1]. For example, imposing Dirichlet boundary conditions for the second-order wave equation using SAT requires decomposing the second derivative operator (sometimes referred to as the "borrowing trick") [19, 33], which complicates the stability
analysis and can lead to less efficient schemes (due to a larger spectral radius of the spatial operator, see e.g. [35] or Paper III). In my view, it is more suitable to use the projection method for these boundary conditions.

The SBP-P method was first presented in [24, 25] as a way of imposing the constraints (the boundary conditions) by incorporating projection operators into the ODE so that only solutions in the subspace where the constraints are fulfilled are considered. Conceptually, the SBP-P method is simpler than the SBP-SAT method since it modifies the ODE the same way regardless of the boundary condition and it does not introduce any free parameters. However, in [13] the SBP-SAT and SBP-P methods were compared with the conclusion that SBP-SAT is preferable since the traditional SBP-P method introduces a zero eigenvalue in the spectrum of the discrete operator. If there is an inconsistency in the initial and boundary data, i.e. if $e_i^\top f \neq 0$ or $d_m^\top f \neq 0$ in (2.10) for example, the error caused by this inconsistency will remain for all time with SBP-P due to the zero eigenvalue, whereas with SBP-SAT it will decrease over time. In [18], an improved projection method was presented where an additional consistent term is added that removes the zero eigenvalue and thus solves this issue of inconsistent data. Additionally, the issues of the zero eigenvalue are minor when solving IBVPs in practice since it is trivial to modify the initial data so that the discrete boundary conditions hold exactly.

Besides SBP-SAT and SBP-P, a stable method for imposing boundary conditions using ghost points has been developed recently [26, 28, 36], but this approach is not considered in this thesis.

Imposing the Neumann boundary condition using SBP-SAT and the Dirichlet boundary conditions using SBP-P, the unconstrained ODE system reads

$$\frac{d^2 v(t)}{dt^2} = c^2 P(D_2 + \tau H^{-1} e_m d_m^\top) P v(t), \quad t > 0,$$

$$v(t) = \hat{f}, \quad t = 0,$$

$$\frac{dv(t)}{dt} = 0, \quad t = 0,$$

(2.11)

where $\tau$ is a scalar parameter tuned for stability, $\hat{f} = P f$ (consistent initial and boundary data), and $P$ is the projection operator given by

$$P = I_m - H^{-1} L^\top (L H^{-1} L^\top)^{-1} L,$$

(2.12)

where $L$ is the linear operator imposing the Dirichlet boundary condition, here $L = e_1^\top$, and $I_m$ is the $m \times m$ identity matrix. The projection operator $P$ is designed to be the orthogonal projection with respect to the inner product defined by $H$, which is equivalent to it satisfying the following self-adjoint property:

$$(u, P v)_H = (P u, v)_H, \quad \forall u, v \in \mathbb{R}^m.$$
Multiplying (2.11) by $P$ from the left and subtracting the result from the original equation results in

$$(I_m - P) \frac{d^2 v(t)}{dt^2} = 0, \quad t > 0,$$  \hspace{1cm} (2.14)

since $P^2 = P$. Integrating in time results in

$$(I_m - P) \frac{dv(t)}{dt} = (I_m - P) dv(0) = 0, \quad t > 0,$$  \hspace{1cm} (2.15)

where the initial data for $\frac{dx}{dt}$ is used in the last step. Integrating in time once again gives

$$(I_m - P) v(t) = (I_m - P) v(0) = \tilde{f} - P \tilde{f} = 0, \quad t > 0,$$  \hspace{1cm} (2.16)

by the definition of $\tilde{f}$. This shows that $v(t) = Pv(t), t > 0$, which is equivalent to $Lv(t) = 0$, i.e. the Dirichlet boundary condition $e_i^\top v(t) = 0$ holds exactly for all time.

The stability of (2.11) is proven using a discrete equivalent to the energy method (used for the continuous problem in (1.2)-(1.7)), where also the parameter $\tau$ is chosen. Multiplying (2.11) by $v_i^\top H$ and using the inner product notation (2.4) leads to

$$(v_t, v_{tt})_H = c^2 (D_1 v_t, D_1 Pv) - e_i^\top D_1 \tilde{v}_i - (D_1 \tilde{v}_i, D_1 \tilde{v})_H - \tilde{v}_i^\top R \tilde{v},$$  \hspace{1cm} (2.17)

where $\tilde{v} = Pv$ and the self-adjoint property (2.13) is used in the last step. Using Definition (2.2.1) to expand $D_2$ leads to

$$(v_t, v_{tt})_H = c^2 ((1 + \tau) e_m^\top D_2 v_{tt} - e_i^\top D_1 \tilde{v}_i - (D_1 \tilde{v}_i, D_1 \tilde{v})_H - \tilde{v}_i^\top R \tilde{v}. $$  \hspace{1cm} (2.18)

Rearranging terms, choosing $\tau = -1$, and using that $e_i^\top \tilde{v}_i = LP v_i = 0$ (by the definition of $P$) results in the energy equation

$$\frac{d}{dt}(\|v_t\|^2_H + c^2(\|D_1 v\|^2_H + v^\top R v)) = 0,$$  \hspace{1cm} (2.19)

where the tilde sign is dropped since $v = \tilde{v} = Pv$, see (2.16). Due to the definition of $R$, the correction term $v^\top R v$ is small and goes to zero as the grid is refined [14]. Therefore the energy estimate (2.19) can be immediately recognized as the semi-discrete analog to the continuous energy estimate (1.6). The equation (2.19) shows that the semi-discrete energy defined by $\|v_t\|^2_H + c^2(\|D_1 v\|^2_H + v^\top R v)$ is constant over time, which is enough to prove stability of (2.11).
3. Temporal discretization

So far only the spatial discretization of IBVPs has been considered. Although the spatial discretization is very important, another crucial aspect of obtaining approximate solutions to IBVPs is the temporal discretization, i.e. how to solve the ODE (2.11). Since the diagonal-norm SBP operators used in the thesis are trivial to invert, and to keep the implementations simple, only explicit time stepping schemes are used in the thesis. For the second-order wave equation (considered in Papers III, IV, and VI) and first-order hyperbolic systems (considered in Paper V), explicit schemes lead to highly efficient implementations since the stability limits (the CFL conditions) for these problems depend linearly on the grid step size $h$. For the IBVPs considered in Papers I and II, the stability limits with explicit time stepping schemes require that the time step scales as $h^2$. For these problems, an argument could be made that implicit solvers are more suitable since the stability restriction with explicit methods can become severe, but closer investigations into this were not necessary for the problems considered in those papers.

For the IBVPs with first derivatives in time considered in the thesis (the incompressible Navier-Stokes equations in Paper I and Maxwell’s equations in Paper V) the standard 4th-order accurate explicit Runge-Kutta scheme (RK4) is used. This is an accurate and efficient method suitable for hyperbolic problems since its stability region includes a portion of the imaginary axis. For the problems with second derivatives in time (the acoustic wave equation in Papers III, IV, and VI and the dynamic beam equation in Paper II) the following two approaches are considered: 1) rewriting the second-order system as a first-order system and solving it using RK4 and 2) using a 4th order accurate explicit two-step finite difference scheme. The main advantage of using RK4 for second-order systems is that the first derivative of the solution is immediately available in the solution vector, which is useful for the shape optimization problem in Paper IV and for treating the dissipative terms in Papers IV and VI. Alternative 2 is preferable for problems where the first derivative is not needed since it requires fewer operations per time step, but it can be problematic if time-dependent source terms are present in the ODE.
4. Accuracy and convergence

As previously mentioned, for diagonal-norm SBP finite difference operators the boundary order of accuracy can be at most $p$ for $2p$th-order accurate interior stencils. This will have an influence on the convergence rate of the overall scheme, i.e. on the rate that the approximate solution converges towards the exact solution when the grid step size is decreased (measured in some norm, in the thesis mostly $H$). If certain conditions are met it can be shown that the convergence rate with diagonal norms is $\min(2p, p + s)$, where $s$ is the order of the highest spatial derivative [6, 32, 34]. Although the proofs in [32, 34] are given for various one-dimensional IBVPs discretized using SBP-SAT, numerical experiments routinely show that the convergence estimate also holds for more complicated problems and problems discretized using other methods (such as multidimensional problems and SBP-P discretizations, for example). For this reason, the convergence rate estimate $\min(2p, p + s)$ is considered a reasonable rule of thumb.

However, the asymptotic convergence behavior is perhaps not the most practically relevant measure of accuracy for numerical methods. For large-scale simulations (for example three dimensional problems) we can typically not afford to perform the simulations with high enough resolution to be in the asymptotic region. So perhaps equally relevant for the accuracy of the results is the so-called error constant. This is the constant factor in the error that is independent of the grid resolution, i.e. the parameter $k$ if the error $\varepsilon$ is given by $\varepsilon = kh^q$, where $h$ is the grid step size and $q$ is the convergence rate. We might be better off using a lower order accurate scheme (small $q$) with a small $k$ rather than a high order accurate scheme (large $q$) with a larger $k$, depending on the grid size $h$ we can use. The idea of decreasing the error constant $k$ is used for the boundary-optimized SBP operators, which are discussed in more detail in Paper I.

Of course, the total error in the final solution will be a combination of the spatial error and the temporal error. But, with the fourth-order accurate time stepping schemes described in Chapter 3, the temporal error is found to be negligible compared to the spatial error for all problems considered in the thesis.
5. Summary of papers

Paper 1
High-order accurate SBP finite difference discretizations of the pressure-velocity formulation of the incompressible Navier-Stokes equations in two spatial dimensions are considered. These equations can for example be used to model the flow of water in pipes or the airflow around buildings. Two new procedures to impose Dirichlet and divergence boundary conditions are derived, one based on the SAT method and the other on the projection method. Additionally, novel boundary-optimized, narrow-stencil, and constant-coefficient second derivative SBP operators are derived and compared to traditional SBP operators. Numerical experiments with an analytical solution show that SBP-SAT and SBP-P are very similar in terms of accuracy and that the new boundary-optimized operators are significantly more efficient than the traditional operators. However, on problems with physical walls (homogeneous Dirichlet boundary conditions), the SBP-P method is found to generate minor vortices close to walls much more efficiently compared to SBP-SAT.

Contributions
The author of this thesis was responsible for performing the numerical experiments and preparing the manuscript of the paper. Editing of the manuscript and developing the numerical methods were done in collaboration between the authors.

Paper 2
The dynamic beam equation and the dynamic Kirchhoff-Love plate equation are biharmonic PDEs used to describe the time-dependent vibrations in beams (1D) and thin plates (2D), respectively. These are PDEs with second derivatives in time and fourth derivatives in space, which makes the boundary treatment of these problems complicated. In a previous work [21], the SBP-SAT finite difference method was used to impose different types of boundary conditions for the dynamic beam equation. In this work, we derive new SBP-P discretizations for those boundary conditions and compare them to the SBP-SAT discretizations. We also consider the equations with piecewise homogeneous coefficients and derive well-posed interface conditions (internal boundary conditions) and discretizations for these problems. Three different methods of
imposing interface conditions for the dynamic beam equation are considered; SBP-SAT, SBP-P, and a new hybrid SBP-P-SAT method. The numerical experiments show that the three methods are mostly similar in terms of accuracy and efficiency. However, minor differences can be found. For example, the SAT method is more accurate for so-called free boundary conditions and the projection method is less stiff when used for interface conditions or so-called clamped boundary conditions (allows for larger time steps). Finally, the dynamic Kirchhoff-Love plate equation with piecewise continuous material coefficients is solved using the hybrid SBP-P-SAT method, to demonstrate that the methods can be extended to two-dimensional problems.

Contributions
The author of this thesis was responsible for preparing the manuscript and partially responsible for the numerical experiments and the conceptualization of the paper. Editing of the manuscript and developing the numerical methods were done in collaboration between the authors.

Paper 3
Using finite differences to simulate wave propagation problems in complex geometries or domains with jumps in the wave speed usually requires splitting the domain into blocks. To obtain a well-posed problem in these situations, specific interface conditions have to be imposed. If, for some reason, the number of grid points of the two blocks at each side of an interface is different (or if they are not located in the same coordinates), interpolation is needed. In a previous work [1], a high-order accurate and provably stable SBP-SAT scheme was derived, including interpolation operators to couple the solutions at both sides of the interface. In this work, the same problem is revisited, and the same interpolation operators are used together with new SBP-P and SBP-P-SAT schemes to impose the interface conditions. Numerical experiments are done showing that the new schemes are similar in terms of accuracy to the SBP-SAT scheme but more efficient since the time step can be decreased significantly (less stiffness).

Contributions
Sole author.

Paper 4
High-order SBP finite differences on curvilinear domains are used in this paper to solve an acoustic shape optimization problem posed as a PDE-constrained optimization problem. The control parameters (the unknowns) are coupled to
the PDE solution by the geometry of the computational domain and are determined by minimizing an objective function using a gradient-based approach. The adjoint framework is used to compute the gradient of the objective function, which can be done efficiently due to the dual consistency properties (or self-adjointness) of the SBP-P-SAT discretization of the acoustic wave equation. The method is used to solve two shape optimization problems of real-world relevance. First, a bathymetry problem where the shape of the seabed is reconstructed from synthetic receiver data, and then an optimal design problem where the shape of the mouth of an air horn is optimized for efficiency (minimizing the amounts of reflections into the horn).

Contributions
The author of this thesis was responsible for performing the numerical experiments and preparing the majority of the manuscript of the paper. Editing of the manuscript and developing the numerical methods were done in collaboration between the authors.

Paper 5
In this paper, multiple new theoretical results with practical relevance for SBP discretizations using the projection method are presented. The core idea is to represent the solution space of the discretization as an inner product space, rather than $\mathbb{R}^N$ equipped with a specific scalar product (defined by $H$). With this approach, the discrete spatial operators are viewed as mappings between vector spaces, which enables certain useful results from linear algebra. Here we apply the concept of pseudoinverses to simplify the construction of SBP-P discretizations. In particular, using the new approach we derive a new projection operator that does not require that the boundary operator has full rank and present a new SBP-P discretization for inhomogeneous problems that does not require the time derivative of the boundary data. As it turns out, an efficient method for imposing continuity at multiblock interfaces can also be described using pseudoinverses. The method presented in the paper (referred to as the embedding method) is similar to the continuous SBP method presented in [8], but here written using operators that map solutions vectors between inner product spaces with and without duplicated points at the interfaces. Numerical experiments measuring the accuracy of the new discretization are done on a first-order hyperbolic system (Maxwell’s equations), showing good agreement with the theoretical expectations.

Contributions
The author of this thesis was responsible for performing the numerical experiments and partially responsible for editing the manuscript.
Paper 6

Although SBP finite differences can be used to discretize domains with somewhat complicated geometries by utilizing domain decompositions and curvilinear mappings, the procedure can become difficult and lead to inefficient schemes for even moderately complex shapes. In these situations, a suitable alternative is to switch to methods based on unstructured grids. In this paper, a collocation spectral element-like method posed in the SBP framework is developed. The focus is on efficient discretizations of the Laplace operator on complex domains decomposed into a large number of quadrilateral elements. The coupling between the elements (imposition of interface conditions) is done using a hybrid SAT and embedding method (as described in Paper V), which results in a discretization where the grid points on the interfaces are not duplicated, i.e. only stored once in the solution vector. Numerical experiments are done showing that the new method is similar in terms of efficiency (accuracy against runtime) to highly efficient boundary-optimized SBP finite differences [29] (for simple enough domains where finite difference methods are reasonable). Additionally, a proof-of-concept simulation is done showing that the method can be paired together with traditional finite difference methods using glue-grid interpolation operators for arbitrarily placed grid points [9] and the SBP-P-SAT method presented in Paper III. With this approach, hybrid discretizations where complex regions of the domain are discretized using the new method and simple regions using traditional finite difference methods are possible, while remaining within the diagonal-norm SBP framework.

Contributions

Sole author.
Det finns ett stort antal fysikaliska processer i världen som kan beskrivas matematiskt med hjälp av partiella differentialekvationer (PDE:er). Två exempel på sådana processer är vågutbredning och strömningsmekanik. Vågutbredningsproblem innefattar bland annat akustiska och elektromagnetiska vågor i atmosfären och under vatten, elastiska vågor i marken och vibrationer i styva balkar och plattor. Strömningsmekaniska processer handlar om hur fluider (vätskor eller gaser) rör sig och påverkas av solida objekt. PDE:erna härleds från fysikens lagar och utgör tillsammans med randvillkor (som specificerar vad som händer på ränderna, till exempel på markytan vid vågutbredning i atmosfären) och initialvillkor (som specificerar systemets tillstånd vid starttiden) så kallade begynnelse- och randvärdesproblem (BRVP). Att kunna lösa BRVPs är ofta mycket användbart inom akademien och industrin. Till exempel är det betydligt snabbare och billigare att matematiskt beräkna luftmotståndet av en ny designad bil än att konstruera bilen och utföra fysiska experiment.


För problemen som behandlas i den här avhandlingen, vågutbredningsproblem och strömningsmekanikproblem, har numeriska metoder med hög noggrannhet visat sig vara effektiva (hög noggrannhet i förhållande till beräkningsstid) [10]. Med låg noggrannhet tenderar vågor som färdas långa avstånd att bli missformade och det är svårt att fånga små strukturer från komplicerade strömningsmekaniska fenomen, som till exempel turbulens.

I den här avhandlingen är huvudfokus på den finita differensmetoden (FDM). FDM är en numerisk metod där beräkningsområden delas upp i ett rutnät och de kontinuerliga operationer från BRVP:n (som derivator och integraler) approximeras diskret på nätet. Noggrannheten avgörs av hur högupplöst rutnätet

Huvudfokus i den här avhandlingen är hur randvillkor ska inkluderas i diskretiseringen så att finita differensoperatorer med SBP egenskaper leder till effektiva och bevisligen stabila diskretiseringar. Som modellproblem behandlas i huvudsak vågekvationen på andra ordningens form, men även diskretiseringar av de inkompressibla Navier-Stokes ekvationerna, den dynamiska balkekvationen och den dynamiska Kirchoff-Love ekvationen presenteras.
Acknowledgements

There are many people that contributed to this thesis being finished. First of all, I want to thank my supervisor Ken. I am very grateful that you gave me the opportunity to pursue a PhD in this field and for your everlasting encouragements. I also want to express my gratitude to all the people I met at TDB, working with and learning from all of you have truly been a pleasure.

I feel very lucky to have ended up doing my PhD in the church of SBP. A huge thanks to Vidar, Martin, Ylva, Jonatan, David and Ken for making a research group feel so much like a group of friends. Vidar, thank you for being a wonderful colleague, a good travel partner, and an exceptional researcher. The projects we did together during these years made my work as a PhD student a lot more interesting and rewarding. Ylva, thank you for being so supportive in the beginning of my PhD and for bringing a much needed positive vibe to the group. Martin, thank you for so often being available to answer my more or less intelligent questions. David, I apologize for the many times you have had to listen to my sometimes incoherent ramblings on things I find important. I am very grateful for our collaboration and I hope we will do many fun and interesting projects in the future. And also, thank you Jonatan (and others) for the tremendous work you put into sbplib, without it Paper VI in particular would have taken much, much longer to finish.

For all the fun moments and for reminding me that there are more important things in life than work, thank you to all my friends in Team LenG. Also, to my closest family, Mamma, Pappa, Oscar, and Victor - thank you all for pushing me to work harder and to be the best version of myself.

Finally, to the love of my life, Lova. Without you I am certain I would not have had the motivation to pursue a PhD, thank you so much for everything you have done for me.
References


A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)