Finite Difference Methods for Wave Dominated Problems

JONATAN WERPERS
Online defence: https://uu-se.zoom.us/j/64178029549

Abstract

Wave models are an important class of models that describe many diverse phenomena such as sound waves, fluid flow, and quantum mechanics. These models are often described mathematically as partial differential equations (PDE). Often these equations do not admit solutions on closed-form and then the only option to study them is numerical methods. These numerical methods must be robust, accurate, and efficient. For spatial discretizations, it is known that higher-order finite-difference methods are efficient, but they often complicate achieving robustness.

In this thesis, we focus on high-order finite-difference methods for solving these wave-dominated PDEs. We use the summation-by-parts (SBP) framework together with simultaneous approximation terms (SAT) for the boundary conditions to prove stability and robustness. This results in efficient numerical methods that are known to converge to the correct solution.

The work in this thesis aims to broaden the scope of these finite-difference methods in different ways. In Paper I and III the framework is extended to two dispersive wave equations, solving challenges arising in both the spatial discretization as well as the time integration. The geometric flexibility of the methods is enhanced for the linear Euler equation and the Schrödinger equation in Paper II and VI by studying both stationary and time-dependent curvilinear grids. Paper IV shows how to use the framework to combine two models to describe and simulate a coupled physical system. In paper V we address a deficiency in the methodology around non-matching grids, showing a way to improve the accuracy and get faster convergence.

Keywords: Finite difference, summation-by-parts, simultaneous approximation term, high-order accuracy, stability, boundary treatment

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

© Jonatan Werpers 2021

ISSN 1651-6214
ISBN 978-91-513-1147-0
urn:nbn:se:uu:diva-435545 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-435545)
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 Wave problems ...................................................................................................... 9
  2.1 Space discretization ......................................................................................... 9
  2.2 Time integration ............................................................................................... 10

3 Summation-by-parts Finite-differences ................................................................. 11
  3.1 Motivating example .......................................................................................... 11
    3.1.1 Well posedness of the continuous problem ................................................. 12
    3.1.2 Semi-discrete problem .............................................................................. 13
    3.1.3 Discrete well-posedness ........................................................................... 14
  3.2 Finite difference methods ................................................................................. 15
    3.2.1 Constructing SBP-FD operators ................................................................. 16
    3.2.2 Composite SBP operators ......................................................................... 18
    3.2.3 Imposing boundary conditions .................................................................. 19
    3.2.4 Accuracy of SBP-SAT schemes ................................................................. 20

4 Summary of Papers .................................................................................................. 21
  4.1 Paper I ............................................................................................................... 21
  4.2 Paper II .............................................................................................................. 21
  4.3 Paper III ............................................................................................................. 22
  4.4 Paper IV ............................................................................................................ 22
  4.5 Paper V ............................................................................................................. 23
  4.6 Paper VI ............................................................................................................. 23

5 Acknowledgements ................................................................................................. 25

6 Sammanfattning på svenska .................................................................................... 26

References .................................................................................................................. 28
1. Introduction

Understanding and predicting wave phenomena might not seem that important to a regular person. While this is probably true for the individual, the level of mastery the scientific and engineering community has of such matters greatly impacts the life of every human being.

Wave phenomena show up in radically separate parts of our lives and underpin many of the technologies that make our modern lives possible. Wireless communication, weather prediction, ultrasound imaging, airplane design, and many others all rely on models that have wave characteristics. Anywhere in the physical world where there is a lack of damping or dissipation of energy, there is a good chance that oscillations will be important and play a big role in understanding the phenomena.

Since the dynamics are naturally described in terms of rate of change in both time and space the models for these phenomena are often expressed as differential equations. When the problem involves both time and space dimensions, as wave phenomena often do, the result of physical modeling is more often than not a partial differential equation (PDE) with initial and boundary conditions. The solutions to these equations then hold the key to understanding the models. They can be used to validate the modeling assumptions by comparing to experimental data, to find areas where the models need refinement. When a suitable model has been found and validated the solutions of these equations are also invaluable for prediction and design tasks.

While posing the differential equation might be relatively straightforward, obtaining the exact solution can be almost arbitrarily hard or even impossible. Therefore when it comes to real-world models, the problem of solving the PDE is often replaced with the problem of solving a different, more approachable, system of equations. This is done using numerical methods and in the ideal case, the method for replacing the PDE is tunable such that the user can make a trade-off between the difficulty of the new problem and the similarity of its solution to the original problem. The way in which this simplification is carried out and how to make it robust is studied in the research area of Numerical Analysis.

The journey from PDE to solution is one paved with mathematics. At every step robustness and efficiency must be considered and mathematically examined so that there can be confidence in the eventual solution. Without efficiency, one might never arrive at the solution and without robustness and confidence, one can not rely on the solution for what you are trying to do.

The issues of stability and robustness are particularly tricky because they require mathematical proofs. In particular, for time-dependent wave equations,
the models are of a nature such that the balance between potential and kinetic energy is exact and there is no dissipation. Therefore the equations live right on the edge of infinite energy growth and any approximation risks nudging the system beyond this boundary. On the other hand, if the approximation adds dissipation the dynamics you are interested in studying might be completely damped in a short time. This places special constraints on numerical methods for wave problems.

Efficiency and robustness are also highly important when it comes to the resource management of researchers. An inefficient method will place undue burden on the computers used to perform the calculation, delay the answers or straight up make the computation infeasible. A method that is not robust will both be harder to develop and use but will also lead to wasted time due to crashed simulations and uncertainty about the source of software bugs. It can even lead to faulty solutions causing bad conclusions to be drawn.

On the other hand, with efficient and robust methods, solvers for new models can be developed without large investment. The models can be iterated and improved. The solutions provided can be trusted. And importantly, solutions can be computed without wasting computer resources.

There are several top-level approaches to discretizing and solving the time-dependent PDE describing wave models. The most popular are finite-volume methods, finite-element methods, discontinuous-Galerkin methods, and finite-difference methods. While they all share many properties there are things differentiating them, making them suitable for different kinds of problems. Finite-volume methods map well to non-linear problems while finite-element methods provide great geometric flexibility and adaptivity. Discontinuous-Galerkin methods bring together high accuracy and geometric flexibility at the cost of computational complexity. Finally, finite-difference methods provide high accuracy and computational efficiency at the cost of flexibility in handling geometry.

This thesis will treat a particular class of finite-difference methods with summation-by-parts properties. The high order of accuracy attainable and the explicit nature of the numerical schemes makes them an excellent match for wave-dominated problems. The summation-by-parts properties ensure that the derived schemes can be made to mimic the energy balance of the physical system, thus leading to proofs of robustness.

The thesis presents a small corner of this field. Through specific applications of physical and numerical modeling, it will advance the state of knowledge of both the systems studied and more generally the numerical methods used to study them. The finite-difference methods are studied and extended to allow more robust computations and improve the framework so that new problems can be accommodated with high efficiency both in terms of computational cost and human effort.
2. Wave problems

One way to view wave phenomena is that they transport some quantity from one place to another. In the case of sound propagation vibration energy is radiating from a sound source as a wave. In quantum physics, the probability density is transported in space as a wave function. Other examples treated in this thesis are waves in elastic ice carrying potential and kinetic energy, or fluids flowing moving mass and pressure.

The fact that this happens in both time and space at the same time, all while preserving mass, energy, or probability places special requirements on the types of methods that we want to use to find solutions. The discretization should inherit the properties of conservation. Often we are interested in waves propagating over a significant distance and a long time, this places further requirements on the methods. It is shown in [9] that to propagate the solution in a hyperbolic problem accurately over long distances we need a method with a high order of accuracy.

An important property for numerical consideration is the speed at which waves travel, and more specifically the phase velocity of the waves. This speed will influence the choice of time integration scheme time-step as discussed in [4]. Of course, the speed will depend on the particular problem and in this thesis we consider equations with phase velocities $\omega \sim \kappa$ and $\omega \sim \kappa^2$, where $\kappa$ is the wavenumber, proportional to the frequency of the wave. Examples of the former are the acoustic wave equation and the linearized Euler equations, and examples of the latter are the Schrödinger equation and the beam equation. Because of the squared dependency on $\kappa$ the waves in the two latter equations travel much faster which complicates the time integration procedures and makes them much more expensive.

2.1 Space discretization

Regardless of the kind of wave problem faced the spatial discretization should respect the properties mentioned earlier. It should have the same conservation properties as the continuous problem and it should have a high order of accuracy to allow efficient propagation of waves over long distances. Finite-difference methods naturally fit this mold. Within the SBP framework they can naturally be constructed to have very high order. In [15], Mattsson et al. presents operators up to order 12. The SBP finite-difference operators also, by construction, mimic the properties of the continuous derivatives that lead
to the conservations properties. Another benefit of their construction is that the application of the difference operator is simply a series of weighted sums. This enables very efficient explicit time-stepping schemes to be used, avoiding solving any linear systems while integrating in time. This is particularly important for problems where $\omega \sim \kappa$, where there is no need for an implicit method in time.

2.2 Time integration

While the error in the solution is often dominated by the space-discretization, the choice of time integration method can have a great impact on the performance of the method as a whole. The wave speed in the particular problem will have a big influence on the choice.

Explicit methods are very efficient per time-step, but the length of that time-step will depend on the discrete phase velocity. The maximum length will be proportional to the speed of the fastest wave on the grid [4], and the frequency of that wave will be proportional to the grid spacing $h$. This means that for problems that have $\omega \sim \kappa$ it is possible to choose the time-step $k \sim h$. For problems with $\omega \sim \kappa^2$, on the other hand, one must pick $k \sim h^2$. Since $h$ is typically small to resolve spatial variations well, the allowed time-step will quickly become vanishingly small. The benefit of each time-step being fast will the drowned by the number of steps needed to reach the end of the simulation.

Implicit methods have a considerably higher cost per time-step, compared to explicit methods. They require solving a system of equations for each time-step. However, they do not have any limitation on the time-step like their explicit counterparts. This means that $k$ will not be tied to $h$ in any other way than the accuracy required in the solution. This trade-off means that for problems where $\omega \sim \kappa$ the best choice is often an explicit method, and for problems with $\omega \sim \kappa^2$, an implicit method.

An interesting version of implicit time-stepping is based on the summation-by-parts framework presented by Nordström and Lundquist [17] and further developed by Boom et al.[2]. These methods are based on finite-differences and the SAT-technique and allow a fully discrete energy estimate if it is available for the continuous problem. This makes them very attractive for wave problems requiring implicit time-stepping. Linders, et al. showed that they are essentially a special case of Runge-Kutta-methods with favorable stability properties[10].
3. Summation-by-parts Finite-differences

Here we will give a short introduction to summation-by-parts operators in general and how they enable a framework for provably stable and robust finite-differences schemes. For more in-depth information the reader is referred to the review papers by Nordström and Svärd [23] and Fernadez et al. [6]. We will start with a motivating example and then present the concept of SBP properties, how the construction of finite-difference operators works, and how to use them to obtain provably stable schemes for initial-boundary value problems.

Of the many different methods that exhibit summation-by-parts properties, the concept originated as a way of making finite-difference operators mimic integration-by-parts from the continuous realm [8]. In [3] the SBP properties were used together with simultaneous-approximation-terms to construct stable finite-difference schemes for hyperbolic systems. While many methods for discretizing PDEs, like FEM, FV, or DG, were not constructed to explicitly have SBP properties, often their stability can be linked to the fact that they do have it. When a discrete operator has the relevant SBP properties it is ensured to behave like its continuous counterpart in important ways.

Finite-difference operators in general lack mechanisms for systematically achieving stability, particularly when the problem involves boundaries and boundary conditions. Deriving stable schemes can be cumbersome and involve ad hoc procedures for each new equation. The usage of SBP properties on the other hand provides a straightforward recipe for implementing boundary conditions in a stable manner.

3.1 Motivating example

The SBP properties can be motivated by studying the continuous setting. We will now investigate an initial boundary value problem and draw parallels between the continuous and semi-discrete setting. The procedure will be very similar for many types of problems and if the continuous problem has the right properties, namely conservation of an energy, there is a very good chance there will be an SBP finite-difference scheme with the same property. The basic steps are the same whether we are studying the linear Euler equations, the Schrödinger equation, or, as we will do here, the wave equation. Many details are left out in the following and for a full treatment of the wave equation the reader is referred to the papers by Virta [24] and Almquist [1].
Consider some domain $\Omega$ with the boundary $\Gamma$ as in Figure 3.1. We can pose the acoustic wave equation

\begin{align*}
  u_{tt} &= \Delta u, \quad t > 0, \ x \in \Omega, \quad (3.1a) \\
  u &= 0, \quad t > 0, \ x \in \Gamma, \quad (3.1b) \\
  u &= u_0(x), \quad t = 0, \ x \in \Omega, \quad (3.1c)
\end{align*}

where (3.1c) is the boundary conditions and (3.1b) is the initial condition. The components of this equation are

- $\Omega$ – The domain in which we want to solve the problem
- $u(x, t)$ – perturbation from equilibrium as a function of time and space
- $\frac{\partial^2}{\partial t^2}$ – second time derivative
- $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$ – laplacian

The properties of these components allow a proof that the equation is well-posed, namely that it has a unique solution.

### 3.1.1 Well posedness of the continuous problem

An important part of proving well-posedness is to show that the solution $u$ is bounded. This is done by taking the inner product of $\frac{\partial u}{\partial t}$ and (3.1).

$$
(u_t, u_{tt})_\Omega = (u_t, \Delta u)_\Omega. \quad (3.2)
$$

The left hand side of this equation can be transformed into $\frac{1}{2} \frac{d}{dt} \|u_t\|^2_\Omega$. And the right hand side

$$
(u_t, \Delta u)_\Omega = -\frac{1}{2} \frac{d}{dt} \|\nabla u\|^2_\Omega + (u_t, \frac{\partial u}{\partial n})_\Gamma. \quad (3.3)
$$

Together with a boundary condition like $u = 0$ or $\frac{\partial u}{\partial n}$, (3.2) becomes

$$
\frac{1}{2} \frac{d}{dt} \left[\|u_t\|^2_\Omega + \|\nabla u\|^2_\Omega\right] = 0. \quad (3.4)
$$
The expression being differentiated on the left can be interpreted as the sum of a kinetic and a potential mathematical energy of the solution. While these are mathematical energies they can in many cases also be mapped to physically meaningful energies.

One of the key properties used here is (3.3) which comes from the integration-by-parts properties of $\Delta$, namely

\[
(u, \Delta v)_\Omega = \int_\Omega u \Delta v \, d\Omega = \\
= \int_\Gamma u \frac{\partial v}{\partial n} \, d\Gamma - \int_\Omega \nabla u \cdot \nabla v \, d\Omega \\
= (u, \frac{\partial v}{\partial n})_\Gamma - (\nabla u, \nabla v)_\Omega.
\]

(3.5)

Here we have seen that using a few properties of the continuous problem we can prove that it is well-posed. One way to derive a numerical scheme for approximating the solution of (3.1) is the come up with discrete counterparts for the components that while approximating the continuous components also preserve the properties used in the proof. Of central importance here is the integration-by-parts property in (3.5). We shall see that by introducing discrete operators that have an analogous property we can arrive at a well-posed discretization.

### 3.1.2 Semi-discrete problem

To arrive at a discrete version of (3.1) the components listed in the previous section will all have to be translated to a discrete counterpart. The way this is done will vary depending on numerical method but we will arrive at

- $\tilde{\Omega}$ – a discrete representation of $\Omega$
- $\tilde{\Gamma}$ – a discrete representation of the boundary of $\Omega$, $\Gamma$.
- $\tilde{u}(t)$ – a vector representing a function on $\tilde{\Gamma}$
- $D_\Delta$ – a matrix approximating the laplacian
- $e^T_{\Gamma} \tilde{u}$, $d^T_{n,\Gamma} \tilde{u}$ – discrete representation of boundary values and normal derivative, respectively.

Using these we can write

\[
\tilde{u}_{tt} = D_\Delta \tilde{u}, \quad t > 0, \quad (3.6a) \\
\tilde{e}_{\Gamma} \tilde{u} = 0, \quad t > 0, \quad (3.6b) \\
\tilde{u} = \tilde{u}_0, \quad t = 0. \quad (3.6c)
\]

We see direct parallels between this formulation and the continuous formulation in (3.1). One wrinkle is the addition of an operator taking $\tilde{u}$ to $\tilde{\Gamma}$. Often it simply picks out the degrees of freedom on the boundary but it can also have more structure including interpolation or extrapolation[5, 13]
### 3.1.3 Discrete well-posedness

To have any hope of a good approximation the discrete problem must also be well-posed. The idea behind the summation-by-parts framework is that if the discrete components satisfy the same kind of properties that were used in the proof of the well-posedness for the continuous problem, then the proof for the discrete problem can mimic the continuous proof.

The first tool we used was the inner product, one for $\Omega$ and one for $\Gamma$, so there will have to be discrete counterparts $\langle \cdot, \cdot \rangle_{\Omega}$ and $\langle \cdot, \cdot \rangle_{\Gamma}$.

To be able to mimic the continuous proof we also need a discrete version of the integration-by-parts property of the Laplacian in these inner products,

\[
\langle \bar{u}, D\Delta \bar{v} \rangle_{\Omega} = \left( e^T \bar{u}, d^T \bar{v} \right)_{\Gamma} - \langle D\nabla \bar{u}, D\nabla \bar{v} \rangle_{\Omega}.
\]

Similarly to how summation can be viewed as a discrete counterpart to integration this property can be viewed as a discrete counterpart to integration-by-parts, i.e summation-by-parts. With this property, we can repeat the same procedure as for the continuous problem. After taking the discrete inner product of (3.6a) with $\bar{u}_t$ we have

\[
\frac{1}{2} \frac{d}{dt} \left[ \|\bar{u}_t\|^2_{\Omega} + \|D\nabla \bar{u}\|^2_{\Omega} \right] = 0.
\]

This gives an energy estimate completely analogous to (3.4) that has the summation-by-parts property we have a well-posed discrete problem.

The method of proving well-posedness in this way is often called the energy method. Given a different problem, with a different PDE and different boundary conditions, if the energy method is applicable with integration-by-parts in the continuous case, there will almost certainly be corresponding discretizations with the summation-by-parts property that yield well-posed discrete problems.

Discretizations of the type (3.6) can be arrived at in many different ways. Perhaps the most common is through a finite-element treatment. Although FEM schemes are rarely presented in this form the fact that the methods work can sometimes be attributed to the that the discrete operators have the summation-by-parts property. In contrast to FEM or FV where the summation-by-parts property is somewhat accidental, finite-difference operators must be constructed explicitly to have it.
While the above is all well and good, a well-posed discrete problem is not necessarily useful. We have not yet touched on the efficiency of the method and well-posedness does not imply stability. The discrete version of the problem (3.6) is a linear system of ODE with the boundary condition acting as constraints. While in this case the constraints are easily eliminated, for other boundary conditions this might not be the case and directly solving the problem on the given form might not be efficient at all. In a following section, we will see how (3.6) is transformed into a different form that in some cases are more applicable.

3.2 Finite difference methods

The basic idea of finite-difference methods is to approximate the derivative of a function $f$ evaluated on a grid $x_i = hi$ using a divided difference. A simple example is the 2nd order central stencil for the first derivative

$$(D_0 f)_i = \frac{1}{2h} (f_{i+1} - f_{i-1}),$$

(3.11)

where $f_i = f(x_i)$. Using Taylor expansion in $h$ one can easily show that

$$(D_0 f)_i = \frac{df}{dx}(x_i) + O(h^2).$$

(3.12)

The order of the approximation refers to how quickly the error approaches zero when $h$ is decreasing. Assuming the function is smooth choosing order of accuracy allows a trade off between accuracy, memory usage, arithmetic operations and other computational resources. An example of a 4th order stencil is

$$(D_{2,4th} f)_i = \frac{1}{12h^2} (-f_{i-2} + 16f_{i-1} - 30f_i + 16f_{i+1} - f_{i+2})$$

$$= \frac{d^2 f}{dx^2}(x_i) + O(h^4).$$

(3.13)

While this stencil take more arithmetic operations to apply that extra cost is offset by allowing a larger $h$ while keeping the error the same and thus allowing less memory and memory bandwidth to be used.

The weights of the stencil can be determined by requiring that the stencil differentiates monomials exactly up to a certain order. Given a $p$th order stencil for the $k$th derivative $D_{k,p}$, we have

$$(D_{k,p} \tilde{x}^j)_i = \tilde{x}^{j-k}_i, \quad \forall j = 0,...,p+k-1,$$

(3.14)

where $\tilde{x}^j = x^j/j!$. Depending on how many terms are allowed there might be one or more solutions to these equations. This fact is used in the boundary stencils of SBP operators to allow enough freedom to achieve the SBP properties sought.
3.2.1 Constructing SBP-FD operators

When using finite-differences for solving boundary value problems one concern is how the stencils close to the boundary are constructed. Since there are no grid points outside the domain the stencils close to the boundary necessarily have to be different than the interior stencils. It turns out that choosing an arbitrary accurate stencil will most likely render the scheme unstable. One way to make a stable method possible is to choose the stencils so that the difference operator has the relevant SBP property. As stated earlier one of the most common SBP properties is one matching IBP of the regular first derivative,

\[ (\bar{u}, D_1 \bar{v})_H = (e_r^\top \bar{u})(e_r^\top \bar{v}) - (e_l^\top \bar{u})(e_l^\top \bar{v}) - (D_1 \bar{u}, \bar{v})_H. \]  

(3.15)

where \( e \) is a vector of zeros with a single element equal to one on the left or right boundary, depending on subscript, and \( H \) is a matrix defining an inner product. Writing this on matrix form and rearranging we can show that choosing the operator

\[ D_1 = H^{-1} \left( Q + \frac{1}{2} e_r e_r^\top - \frac{1}{2} e_l e_l^\top \right), \quad Q + Q^\top = 0, \quad H = H^\top > 0, \]  

(3.16)

will satisfy (3.15). So to find a useful operator we can take \( H \) and \( Q \) as unknown and impose the accuracy conditions (3.14) on them. Here \( Q \) ensures that the skew-symmetric nature of the first derivative is shared by this discrete operator and, \( H \) will define the inner product in which the SBP property holds. The interior parts of the operator, shown in blue in Figure 3.2 are given by the inner stencil. The red parts however will be the solution of a linear system of accuracy equations

\[ Q \bar{x}^j + H \bar{x}^{j-1} = \left( \frac{1}{2} e_l e_l^\top - \frac{1}{2} e_r e_r^\top \right) \bar{x}^j \quad j = 0, \ldots, q. \]  

(3.17)

where \( q \) is the order of accuracy for the boundary stencils. In Figure 3.2 the structure of a operator with an interior order of accuracy of 4 is shown. Notice that even though the interior stencil would fit at row 3 and 4, in order for a solution to (3.17) to exist they must be different. Figure 3.3 shows the structure of the corresponding \( Q \) where blue elements are induced by the interior stencil and the the red elements are the unknowns solved for in (3.17).

The exact number of boundary stencils can vary depending on the SBP-property, the order of accuracy, the order of derivative, inner stencil, and so on. Often the system of equations is under-determined which leaves some parameters to be tuned, often to minimize some weighted sum of the truncation errors at the boundary. Mattsson discusses these issues for the second derivative in [11] and for the third and fourth derivative in [12].

The order of accuracy for the boundary stencils will be impacted by the inner stencil and the type of inner product chosen. When using a diagonal \( H \)
Figure 3.2. Sparsity patterns of SBP matrices with boundary regions colored red.

Figure 3.3. Sparsity patterns of the skew symmetric part of $D_1$ with boundary regions colored red.

as in Figure 3.2 $q$ will be capped at $p/2$. If $H$ is allowed to be non-diagonal in the boundary regions it is possible to find boundary stencils of order $p - 1$. However, this comes with complications regarding variable coefficients and non-cartesian grids which renders these operators less used. Non-diagonal $H$ can also be constructed from implicit finite-differences, see [3], giving excellent dispersion properties at the cost of solving a linear system to apply the operator.
3.2.2 Composite SBP operators

Most commonly SBP finite-difference operators are constructed for a single dimension. The most common are the first and second derivative satisfying

\[
(\bar{u}, D_1 \bar{v})_H = (e^T r \bar{u}) (e^T_1 \bar{v}) - (e^T_1 \bar{u}) (e^T_1 \bar{v}) - (D_1 \bar{u}, D_1 \bar{v})_H, \tag{3.18}
\]

\[
(\bar{u}, D_2 \bar{v})_H = (e^T r \bar{u}) (d^T_{1,2} \bar{v}) - (e^T_1 \bar{u}) (d^T_{1,2} \bar{v}) - (\bar{u}, \bar{v})_M, \tag{3.19}
\]

for different orders [21] and grids [14, 15, 2]. While it is possible to directly construct operators for other SBP properties, like in [7] it is also possible to combine these basic operators to build new ones with different SBP properties. When this route is available it is considerably less work.

One very important case of this is the construction of SBP operators for higher dimensions than one. A simple example is the Laplace operator in two dimensions. Using the Kronecker product we can define

\[
D_\Delta = D_2 \otimes I + I \otimes D_2,
\]

\[
H_{2D} = H \otimes H,
\]

\[
M_{2D} = \text{diag}(M \otimes I, I \otimes M),
\]

\[
H_\Gamma = \text{diag}(H,H,H,H),
\]

\[
e_\Gamma^T = \begin{bmatrix} e^T_1 \otimes I & e^T_1 \otimes I & I \otimes e^T_1 & I \otimes e^T_1 \end{bmatrix},
\]

\[
d^T_{1,\Gamma} = \begin{bmatrix} d^T_{1,1} \otimes I & d^T_{1,r} \otimes I & I \otimes d^T_{1,1} & I \otimes d^T_{1,r} \end{bmatrix}.
\]

We then have an operator that satisfies

\[
(\bar{u}, D_\Delta \bar{v})_{H_{2D}} = \left( e^T_\Gamma \bar{u}, d^T_{1,\Gamma} \bar{v} \right)_{H_\Gamma} - (\bar{u}, \bar{v})_{M_{2D}}, \tag{3.21}
\]

which can be used to solve for example the wave equation in two dimensions.

Another example is \(D_1^2\) that satisfies

\[
(\bar{u}, D_1^2 \bar{v})_H = (e^T_1 \bar{u}) (e^T_1 D_1 \bar{v}) - (e^T_1 \bar{u}) (e^T_1 D_1 \bar{v}) - (\bar{u}, \bar{v})_H, \tag{3.22}
\]

which is similar to (3.19) but more closely matches the corresponding IBP property of \(\frac{\partial^2}{\partial x^2}\), and may be applicable for some problem. Another example is the IBP property needed for the variable advection equation,

\[
u_t + a(x) u_x = 0, \quad x \in (0,1), \tag{3.23}
\]

which using \((u,a(x)v_x) = auv\big|_0^1 - (a'(x)u,v) - (a(x)u_x,v)\) can be shown to be well behaved. A difference operator for \(a(x)\frac{\partial}{\partial x}\) with a corresponding SBP property can be written as

\[
D_{1,a} = \frac{1}{2} \left[ AD_1 + D_1 A - A_x \right], \tag{3.24}
\]

where \(A\) and \(A_x\) are a diagonal matrices with \(\bar{a}\) and \(\bar{a}_x\) on the diagonal. Together with curvilinear multi-block grids, the 1D operators can be combined to yield SBP operators for arbitrary domains in both 2D and 3D.
3.2.3 Imposing boundary conditions

Having the difference operator is only the first step towards having a working numerical method. If we for example are interested in solving the scalar one dimensional advection equation

\[ u_t + u_x = 0, \quad x \in (0, 1), \quad t > 0, \]
\[ u = g(t), \quad x = 0, \quad t > 0, \tag{3.25} \]

which admits the energy estimate

\[ \frac{d}{dt} \|u\|^2 = g^2 - u_r^2. \tag{3.26} \]

We may use the SBP operator from the previous section to write a semi-discrete equation

\[ \bar{u}_t + D_1 \bar{u} = 0, \quad t > 0, \]
\[ e_1^\top \bar{u} = g(t), \quad t > 0. \tag{3.27} \]

In this particular case, because of the simplicity of \( e_1 \) the first element of \( \bar{u} \) can simply be eliminated and the discrete energy estimate will be

\[ \frac{d}{dt} \|\bar{u}\|_H^2 = g^2 - (e_1^\top \bar{u})^2. \tag{3.28} \]

Often however we are not so lucky. In the case of systems of PDE or equations involving higher-order derivatives simply eliminating the discrete boundary conditions will not be possible. To efficiently impose the boundary condition several different methods have been developed. Using a simultaneous approximation term (SAT) is perhaps the most common in association with SBP-FD but there are also others like ghost points[20] and the projection method [18, 19, 16].

Imposing the boundary condition using SAT involves adding a penalty term that aims to drag the solution towards the correct value,

\[ \bar{u}_t + D_1 \bar{u} = \tau (e_1^\top \bar{u} - g(t)) \quad t > 0. \tag{3.29} \]

Since the true solution satisfies the boundary condition this does not affect the consistency of the scheme. However, the parameter \( \tau \) will affect the stability and must be chosen so that stability can be proved. Choosing \( \tau = -H^{-1}e_1 \) yields

\[ \frac{d}{dt} \|\bar{u}\|_H^2 = g^2 - (e_1^\top \bar{u})^2 - (e_1^\top \bar{u} - g)^2, \tag{3.30} \]

which is the same as (3.28) with an extra damping term.

Similar techniques work for many problems and even for other methods like FEM and DG. If the continuous problem admits an energy estimate and the discretization has the correct SBP property the SAT method can be applied in the same way. It also works for implementing interface conditions as is done in Paper V. While the method is very flexible, determining the penalty parameter \( \tau \) can sometimes be an involved affair.
3.2.4 Accuracy of SBP-SAT schemes

Given a scheme based on a $2p$th order accurate SBP operator one might expect the solution to converge to with the order $2p$. Unfortunately, this is not always the case. The boundary stencils in an SBP operator based on a diagonal inner product are restricted to be only $p$th order accurate. So then one might expect the order of accuracy for the full scheme to be $p$. Luckily, this is not the case either. From a discrete energy estimate, it is possible to show that the accuracy will be at least $p + \frac{1}{2}$. Svärd and Nordström showed that for equations involving first derivatives we can expect order $p + 1$ [22]. Even better, for equations with second derivatives, Wang and Kriess showed that order $p + 2$ is expected [25]. The picture is further complicated by 4th derivative operators which are constructed to have a boundary accuracy of $p - 2$.

How to show the exact order of accuracy for an SBP scheme is not fully understood. Nevertheless, a wide range of numerical experiments for many different problems and setups show that with an interior order of $q$, a boundary order of $p$, and a problem with a $k$th derivative one can expect an order of accuracy of $\min(q, p + k)$. Table 3.1 shows the expected order of convergence for a few operators and derivatives.

<table>
<thead>
<tr>
<th>Name (Interior)</th>
<th>Boundary</th>
<th>$\partial_x$</th>
<th>$\partial_{xx}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd</td>
<td>1st</td>
<td>2nd</td>
<td>2nd</td>
</tr>
<tr>
<td>4th</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
</tr>
<tr>
<td>6th</td>
<td>3rd</td>
<td>4th</td>
<td>5th</td>
</tr>
<tr>
<td>8th</td>
<td>4th</td>
<td>5th</td>
<td>6th</td>
</tr>
</tbody>
</table>

Table 3.1. The expected order of accuracy for a few combinations of operators and problems.
4. Summary of Papers

4.1 Paper I

Paper I analyses a dispersive non-linear model of nerve-signal propagation. The model is an alternative to the more classical Hodgkin-Huxley model of signal propagation. The equation contains an anti-dispersive non-linear second derivative and a dispersive fourth derivative. The contradicting nature of the two terms gives rise to soliton solutions.

The numerical method derived is based on the SBP-SAT framework and new finite-difference operators of different orders are constructed that allow a joint treatment of the two terms. Two different boundary conditions are analyzed and implemented using penalty terms. An explicit time integration method is derived based on centered finite-differences. The fully discrete method is shown to be linearly stable.

The method is validated by comparing numerical approximations to an analytical solution. The robustness of the method is queried by studying non-smooth initial conditions and long-running simulations. Finally, the efficiency of the different orders is examined and higher-order spatial discretizations are shown to be advantageous.

Contributions

J. Werpers developed the numerical methods in collaboration with K. Mattsson. J. Werpers implemented the methods, conducted and designed the numerical experiments. The manuscript was written in collaboration between the two authors.

4.2 Paper II

Paper II presents two finite-difference schemes for the linearized Euler equations with the aim of computing sound propagation in a non-homogeneous atmosphere. One scheme is based on standard central differences and one on upwind stencils. The SBP-SAT framework is used to provide high-order, provably stable discretizations on curvilinear grids.

The upwind operators have high-order upwind stencils that introduce a small amount of dissipation. For a simplified 1D case the interaction of these operators with point source discretizations is compared to the centered operators. It is demonstrated that the discretization of the point source can be
simplified when using the upwind operators by omitting the smoothness conditions typically used. The dissipation ensures that the high-frequency error is damped and full order convergence is observed.

The analysis is validated with numerical experiments and convergence tests for different setups, including a variable topography, variable air density, and windy conditions.

**Contributions**

J. Werpers was heavily involved in the development of the ideas and the direction of investigation. He helped in the design of numerical experiments. The manuscript was prepared in collaboration between the three authors.

### 4.3 Paper III

A model for simulation of acoustic and flexural-gravity waves in ice-covered oceans is presented. The equations describe the dynamics and interactions of ocean waves when they interact with a thin layer of floating elastic ice. The paper develops a high-order finite-difference scheme using the SBP-SAT methodology for solving these equations.

The method couples two domains of the acoustic wave equation using a numerical interface. The beam equation, modeling the ice, is coupled to the boundary of the acoustic wave equation. To handle the stiff nature of the equations an 8th order SBP scheme is used for time integration.

Stability is proven and higher order of convergence is shown in numerical experiments. Finally, some physical properties of the system are investigated using the numerical method.

**Contributions**

The model was developed and analyzed in collaboration between the three authors. J. Werpers developed, analyzed, and implemented the numerical method with input from the other authors. J. Werpers designed and performed the numerical experiments. The manuscript was prepared in collaboration between the three authors.

### 4.4 Paper IV

The paper develops and tests a model for the discharge of a seismic air-gun. The model describes how, when the air-gun fires, the air is ejected into the water surrounding the air-gun and how the formed bubble interacts with the air-gun. The aim is to understand how the design and use of the air-gun influences the frequencies of sound produced.
As opposed to earlier attempts using lumped parameter models the presented model resolves the interior dynamics of the air-gun using the 1D Euler equation. The SBP-SAT framework is used to derive a robust coupling between the PDE model of the air-gun and the ODE model of the bubble.

Contributions
The authors jointly developed the model and coupling conditions. The numerical methods were implemented and described in the manuscript by J. Werpers.

4.5 Paper V
This paper improves on methods for handling interfaces with non-conforming grids within the SBP-SAT framework. Previous methods are forced to sacrifice one order of accuracy to prove stability. In this paper, a method is presented for problems involving the Laplacian operator that maintains the accuracy of conforming interfaces while still allowing a proof of stability.

The improved accuracy is achieved by introducing a new set of interpolation operators that have a lower truncation error. We prove that these operators will always exist for diagonal inner products regardless of the grid point distributions. The operators are used to derive non-conforming grid interfaces for the wave equation, the heat equation, and the Schrödinger equation. The schemes all achieve that same order of accuracy as the conforming case and are provably stable.

The schemes are implemented and compared to a previous method, demonstrating the gain in accuracy.

Contributions
The authors developed the ideas and proofs in collaboration. J. Werpers provided parts of the implementation of the numerical methods and helped design and run the numerical experiments. The manuscript was prepared in collaboration between the three authors.

4.6 Paper VI
Paper VI develops a numerical scheme for the Schrödinger equation on deforming domains. This model is of interest when studying quantum billiards and chaos. The scheme is based on summation-by-parts finite-differences and weakly imposed boundary conditions. The deforming domain is handled using a time-dependent mapping from a static logical grid.

The derived spatial discretization is shown to be high order and provably stable. It is also shown to preserve physical properties of the continuous model, e.g conservation of probability. The resulting coefficient matrix is
time-dependent which complicates the time integration, but we prove that the Mangus-Lanczos method is applicable.

We implement the method and validate the properties through numerical experiments. We show that the scheme converges as expected. Finally, reproduce a physical experiment showing the occurrence of a berry phase for a quantum-mechanical system.

Contributions
J. Werpers was heavily involved in the development and analysis of the numerical method. He also assisted in organizing and writing the manuscript.
5. Acknowledgements

To begin I would like to thank my advisor Ken Mattsson for convincing me to apply for the Ph.D. position that has led to this thesis. The things I have learned during these past years have been some of the most fun and interesting things I have ever encountered. I would also like to thank Ken for all the support, guidance, and all the opportunities he has provided me. Secondly, I want to thank my co-advisor Erik Sjöqvist for the many fun discussions we have had. Also, I am grateful to my collaborators Eric Dunham, Leighton Watson, and Siyan Wang for their friendliness, and their willingness to share their ideas and discuss.

Next, I want to send a huge thank you to all my present and former colleagues at TDB for providing a fun and engaging work environment. Thank you for all the laughs and for your willingness to answer all my silly questions. I remember talking to some of you about how TDB is like a mini version of a complete university with experts in mathematics, physics, biology, geophysics, and many other things. I will certainly miss having such easy access to so much knowledge. Special thanks to some of my former colleagues, Fredrik Hellman for the many interesting discussions and for teaching me your whistling techniques, and Timofey Mukha for all the fun League of Legends games and for expanding my Russian vocabulary.

Second to last I would like to thank my closest colleagues Martin, Ylva, and Vidar. I am very grateful for having you as my friends. Martin, a special thanks to you for being the best scientific sparring partner imaginable.

Till sist vill jag tacka min familj Karin, Olof, Bäbo II, Felix och mamma och pappa för allt stöd och för att ni fyller mitt liv med så många mysiga, roliga och avkopplande stunder.
Att förstå och förutse vågfenomen kan för gemene man verka oviktigt. Antagligen stämmer för den enskilde individen, men mänsklighetens samlade kunskap kring sådana ämnen har en otroligt stor inverkan på hur vi lever våra liv.

Vågfenomen dyker upp i vitt skilda delar av vår verklighet och ligger till grund för många av de teknologier som gör våra moderna liv möjliga. Trådlös kommunikation, väderprognoser, ultraljud, flygplansdesign och många andra saker är allt sådant som i grunden bygger på olika typer av vågutbredning. När ett fysikaliskt system har låg inbyggd dämpning är det troligt att oscillationer och vågor kommer att spela en stor roll i förståelsen av systemet.

Eftersom dynamiken naturligt uttrycks i termer av förändringar i tid och rum så beskrivs ofta de matematiska modellerna i termer av differentialekvationer. När problemet involverar både rums- och tidsdimensioner är resultatet av den fysikaliska modelleringen allt som oftast en partiell differentialekvation (PDE) med begynnelsevillkor och randvillkor. Lösningarna till dessa ekvationer ligger till grund för att förstå det ursprungliga fenomenet. De kan, genom att jämföras med experiment, användas för att validera modelleringsanslaganden och för att hitta områden där modellen behöver förbättras. När man har en validerad modell blir lösningarna ovärdelig verktyg för prediktion och design.


Vägen från PDE till lösning kantas av matematik. Vid varje steg måste vikt läggas vid att bevara robusthet och effektivitet så att det går att lita på den slutliga lösningen. Utan effektivitet löper man risken att aldrig kunna räkna så länge att man kommer fram till lösningen, även om det i teorin skulle gå. Om metoden inte är robust kommer vi inte kunna lita på att vi har funnit rätt lösning.

Stabilitet och robusthet är extra känsliga egenskaper som kräver matematiska bevis. Speciellt när det gäller tidsberoende vågekvationer, är modellerna

Egenskaperna effektivitet och robusthet är också väldigt viktiga när det kommer till att hushålla med forskningsresurser. En ineffektiv metod kommer att slösa beräkningsresurser och energi när beräkningarna utförs, försena svaren och kan till och med göra beräkningen omöjlig att slutföra. En metod som inte är robust kommer både att vara svårare att utveckla och att använda. Dessutom kommer den leda till slösad tid för forskaren i och med kraschade simuleringar och osäkerhet i var eventuella mjukvarufel kommer ifrån. I värsta fall leder det till felaktiga lösningar och i förlängningen felaktiga vetenskapliga slutsatser.


References


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2019

Editor: The Dean of the Faculty of Science and Technology

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”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-435545