Freie Universität Berlin

A Multigrid-based Multiscale Numerical Scheme for Shallow Water Flows at Low Froude Number

vorgelegt von

Stefan Vater

Berlin, 2013

A Multigrid-based Multiscale Numerical Scheme for Shallow Water Flows at Low Froude Number

Inauguraldissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich Mathematik und Informatik
der Freien Universität Berlin

vorgelegt von

Stefan Vater

Erstgutachter & Betreuer:
Prof. Dr. Rupert Klein, *Institut für Mathematik,*
Freie Universität Berlin

Zweitgutachter:
Prof. Dr. Michael L. Minion, *Institute for Computational and*
Mathematical Engineering (ICME), Stanford University

Tag der Disputation:
19. Dezember 2012

Berlin, 2013

Abstract

A new multiscale semi-implicit scheme for the computation of low Froude number shallow water flows is presented. Motivated by the needs of atmospheric flow applications, it aims to minimize dispersion and amplitude errors in the computation of long-wave gravity waves. While it correctly balances "slaved" dynamics of short-wave solution components induced by slow forcing, the method eliminates freely propagating compressible short-wave modes, which are under-resolved in time. This is achieved through a multilevel approach borrowing ideas from multigrid schemes for elliptic equations. The scheme is second-order accurate and admits time steps depending essentially on the flow velocity. It incorporates a predictor step using a Godunov-type method for hyperbolic conservation laws and two elliptic corrections for the computation of the numerical fluxes.

First, a multilevel method is derived for the one-dimensional linearized shallow water equations. Scale-wise decomposition of the data enables a scale-dependent blending of time integrators with different principal features. To guide the selection of these integrators, the discrete-dispersion relations of some standard second-order schemes are analyzed, and their response to high-wave-number low-frequency source terms is discussed. In particular, the implicit trapezoidal rule and two backward differentiation formula (BDF) schemes are considered. The resulting method essentially consists of the solution of a Helmholtz problem on the original fine grid, where the differencing operator and the right hand side incorporate the multiscale information of the discretization. The performance of the new multilevel method is illustrated on a test case with "multiscale" initial data and a problem with a slowly varying high-wave-number source term.

The scheme for simulating fully nonlinear shallow water flows is a generalization of a projection method for the zero Froude number equations. Therefore, the method described in Vater and Klein (*Numer. Math.* 113, pp. 123–161, 2009) is extended to account for time dependent bottom topography. Numerical simulations show that the method is well-balanced in that it can reproduce the steady state of a lake at rest with non-trivial bottom topography. Furthermore, the results of two other test cases verify the correct representation of time dependent bottom topography.

Finally, the multiscale semi-implicit method for the nonlinear shallow water equations at low Froude numbers is derived. This incorporates two different extensions of the aforementioned projection method by incorporating the local time derivatives

of the height. These extensions are based on an implicit trapezoidal and a BDF(2) time discretization, respectively, for the linear implicit part of the scheme. The two variants are combined in the multiscale method, which incorporates the solution of a multilevel Helmholtz problem, similar to the linear case. The method is implemented in one space dimension. The convergence of the method is analyzed by means of different test cases. Moreover, the method's balancing properties are addressed for the lake at rest.

For the derivation of the method, the shallow water model is extended to include time-dependent bottom topography. This is needed to simulate a source term which directly acts on the local divergence of the flow, such as latent heat release due to local condensation in atmospheric flow problems. Furthermore, the asymptotic regime of fast gravity waves traveling over short-range topography is considered, which is equivalent to the low Mach number regime of long-wave acoustic waves interacting with slow advection as described in Klein (*J. Comput. Phys.* 121, pp. 213–237, 1995). The numerical results of the proposed method suggest that the scheme correctly reproduces this regime, and can be therefore considered as a so-called asymptotically adaptive numerical method.

Acknowledgments

First of all, I would like to express my gratitude to Prof. Dr. Rupert Klein from Freie Universität (FU) Berlin for continuous support, fruitful discussions and critical suggestions on how to improve my studies. Under his supervision, I got into contact with numerous internationally renowned scientists and had the opportunity to present my work at different international conferences. Furthermore, I would like to thank Dr. Michael Minion from Stanford University, who agreed to be the co-referee and provided many helpful comments during the finalization of this work.

At FU Berlin and the Potsdam Institute for Climate Impact Research (PIK) I experienced a stimulating working atmosphere and interesting discussions with all group members. I am particularly indebted to Dr. N. Botta, who patiently answered my countless questions and introduced me to his scientific software for the numerical implementation of my algorithms in the beginning of my thesis. Daniel Ruprecht and Carsten Scharfenberg were always open to discussing with me several detailed questions, which arose during my work.

Moreover, I would like to thank Prof. Dr. Jörn Behrens from University of Hamburg, who gave me the opportunity to start my work as a PostDoc in his group before I finished this thesis.

I would like to express my gratitude to my wife Theresia Petrow and my family. They supported me during all these years of research and shared with me good and difficult times. Without their encouragement, this work would have probably never been finished. Last, but not least, I am indebted to Felix Rieper and Martin Weiser for revising large parts of the manuscript. Parts of this work were published in Vater *et al.* (2011). The two anonymous referees of this article provided very constructive reviews that also helped to improve this thesis.

This work benefited greatly from free software products. Without these tools – such as LaTeX, the GNU C/C++ compiler and the linux operating system – a lot of tasks would not have been so easy to realize. It is my pleasure to thank all developers for their excellent products.

Berlin, September 2012   *Stefan Vater*

Contents

1 Introduction

The overwhelming complexity of the earth's atmospheric system involves a variety of processes, which act on different space and time scales. These scales range over several magnitudes from microphysics to planetary scale phenomena. Examples of important processes are the balances governing relevant meteorological regimes such as the hydrostatic equilibrium, which states the balance between the forces arising from gravitation and the vertical pressure gradient. Furthermore, phenomena present on different scales interact with each other through the nonlinear behavior of the system. This could be the small scale flow divergence induced by the energy budget of local cloud processes, which modulates larger scale quasi-linear wave phenomena. Disregarding these peculiarities potentially leads to erroneous results in the computation of atmospheric flow problems arising from meteorology, and advanced numerical techniques are required for the solution of such problems. The correct representation of the above mentioned balances, for example, is achieved by using so-called well-balanced numerical methods (see Botta *et al.*, 2004, and references therein). These schemes ensure that an initially exact balance is preserved and numerical errors do not accumulate over time leading to unphysical solutions.

Concerning the topic of time integration, it must be further regarded that the governing equations, which are the compressible flow equations applied to the rotating earth, admit wavelike solutions as perturbations of the above mentioned balances. These waves are associated with compressibility (acoustic waves), buoyancy (gravity waves) and inertia (Rossby waves), and act on well separated space and time scales (Klein, 2010). While the fast acoustic (or sound) waves are usually of little interest in meteorological applications, the much slower synoptic and planetary wave patterns associated with inertia and advection are most relevant for predicting the weather. However, for stability reasons the time step in explicit numerical methods is dictated by the speed of the fastest spread of information in a given problem, which is expressed

by the *Courant-Friedrich-Lewy* (CFL) condition (Courant *et al.*, 1928). This often leads to a too severe time step restriction for practical computations.

There have been basically two different approaches to deal with this problem. One possibility is to replace the governing equations by reduced model equation sets, which do not include the fast waves, before the construction of a numerical scheme. The most notable are the *anelastic* or *quasi-Boussinesq* (Ogura and Phillips, 1962; Lipps and Hemler, 1982; Bannon, 1996) and the *pseudo-incompressible* approximations (Durran, 1989), which "filter out" acoustic waves. The second approach concerns the numerical discretization, and aims at dealing with the fast components of the flow in a stable manner (see Durran, 2010, for an overview). This is commonly achieved by an approximation which "splits" the fully compressible equations into the fast and slow wave components, where the problem corresponding to the fast waves is linear and thus easier to solve. The fast part is either discretized by an explicit method using a much smaller time step than for the slow solution part (so-called *split-explicit* or *multirate* methods), or it can be treated implicitly, leading to *semi-implicit* methods.

At a first glance the application of filtered equation sets may seem to be attractive, since the numerical treatment of the disparate scales can be avoided. In fact, such approximate "sound-proof" model equations, which are the analogues of the classical incompressible flow equations, are commonly used in current computational simulations of small-scale atmospheric processes, such as cloud formation, local storms, or pollutant transport on city-scales (Lipps and Hemler, 1982; Durran, 1989; Grabowski, 1998). However, these models are usually obtained by *single-scale asymptotic analyses* or scaling arguments that only capture the essence of the processes specific to the considered scale. Their applicability to realistic regimes of stratification is still an active research topic (see, e.g. Klein *et al.*, 2010).

The numerical treatment of the *fast wave–slow wave* problem benefits from the fact that this approach aims for a consistent approximation of the fully compressible equations. Also, the splitting ansatz tries to accommodate to some extent for the multiscale behavior of the atmospheric system. However, this procedure has also its weaknesses. By adopting implicit time discretizations for the fast components of the flow, one may overcome the time step limitations, but is then faced with a potentially undesirable numerical dispersion behavior of the resulting numerical schemes: most – if not all – existing implicit schemes of second or higher order of

convergence achieve large time step stability by slowing down the short-wave acoustic modes. At the same time, some quite popular schemes essentially preserve their amplitude, which may seem desirable at a first glance, but can result in a source of nonlinear instabilities in practice. To see this, consider a full-fledged atmospheric flow model which incorporates parametrization of subgrid scale diabatic effects, such as latent heat release from localized condensation. Local, small-scale heat release will set up non-zero flow divergences, and these are projected partially onto slow fake numerical acoustic modes. These modes will not disappear subsequently unless removed artificially, and they can influence the further flow evolution by interacting themselves with various subgrid scale parametrizations. A simple solution is to employ a first-order implicit discretization, such as the implicit Euler scheme. This removes the short-wave fake acoustic modes, but it will also dampen or modify the long-wave modes in an undesired fashion.

The need for numerical methods that capture the multiscale behavior of the atmosphere can be seen from the following example. The so-called hydrostatic primitive equations (HPEs) are widely used in current operational general circulation models (GCMs) for global atmospheric flow simulations. They are obtained from the full three-dimensional compressible flow equations in the limit of large horizontal-to-vertical scale ratios. While this asymptotic limit suppresses vertically propagating sound waves, it does support long-wave horizontally traveling acoustic modes, i.e., Lamb waves. These Lamb waves are sometimes considered non-negligible for planetary-scale dynamics. Furthermore, there are indications that effects of compressibility affect large-scale, deep internal wave modes of the atmosphere in a non-trivial fashion (Davies *et al.*, 2003). Therefore, approximate "sound-proof" model equations may not provide an accurate representation of planetary scale flows. On the other hand, modern high-performance computing hardware is beginning to allow modelers to use grids with horizontal spacing in the range of merely a few kilometers even for planetary-scale simulations (see e.g., Ohfuchi *et al.*, 2004). At such high resolution, the hydrostatic approximation breaks down, and one enters the scale range of sound-proof model applications. Although these issues are still under debate (Smolarkiewicz and Dörnbrack, 2008), keeping effects of compressibility in planetary-scale simulations seems desirable, and the challenge arises from combining large-scale compressible flow representations with essentially sound-proof modeling of the small scale dynamics.

For the mathematical study of interactions across scales, techniques from *multiple-scales asymptotics* (Kevorkian and Cole, 1996) are increasingly used. These are extensions of the classical single-scale asymptotic method (also known as regular perturbation analysis). In the latter one, a small non-dimensional parameter of the problem and a special (asymptotic) expansion of the dependent variables are employed to obtain simplified equation sets, which still account for the physical effects characteristic to the specific scale. In multiple-scales asymptotic analysis the asymptotic expansion is generalized in that the variables artificially depend on more than one space or time scale. This enables the study of effects arising across scales. Since the asymptotic analysis directly relates a reduced model to the full compressible flow equations, it is a natural starting point for the development of numerical methods applicable to the considered singular regimes (Klein, 2004). In this context, the notion of *asymptotically adaptive numerical methods* (Klein, 2000; Klein *et al.*, 2001) was created. Such schemes should be robust, uniformly accurate, and efficient in the vicinity of certain asymptotic regimes and over a variety of relevant applications. The idea is closely related to so-called *asymptotic preserving* (AP) schemes (see Cordier *et al.*, 2012, and references therein).

The aim of this thesis is therefore to develop such an asymptotically adaptive numerical method, which is able to correctly simulate large scale compressible flow phenomena with high resolution. In this initial attempt not the full dynamics of the atmosphere are considered. Instead, this work deals with the *shallow water equations*, which describe the vertically averaged motion of an incompressible fluid with a free surface. By "shallow" one refers to the small aspect ratio between the vertical depth and a typical horizontal length scale of the problem, which justifies the hydrostatic assumption, i.e. the pressure balances the weight of the fluid. However, these equations are not only a good model for representing river flow or large scale oceanic motions (such as tsunamis). While ignoring the presence of stratification, the shallow water equations incorporate the effects of gravity and can account for the earth's rotation and for bottom topography by the addition of appropriate source terms. Therefore, they are prototypical of the hydrostatic primitive equations and are often used in the development of numerical methods for atmospheric flow problems.

Due to the vertical averaging of the prognostic variables, the shallow water equations only admit external waves. However, the external gravity waves in shallow water flows

are the equivalent to Lamb waves in the compressible flow equations (Kalnay, 2003). The considered asymptotic regime consists of long-wave acoustic waves (Lamb waves) interacting with slow advection. This is equivalent to the regime of fast gravity waves moving over short range topography in the shallow water context. The additional effect of small scale flow divergence induced by local diabatic effects is modeled by a time dependent bottom topography. In this context, the shallow water model represents one of the most challenging parts in the development of numerical methods for the simulation of planetary scale atmospheric flows at high resolution.

The analysis of the regime of fast gravity waves moving over short range topography reveals that it essentially consists of long-wave linearized shallow water flow interacting with small scale flow balancing the influence of the rough topography (see Chapter 2 for details). Therefore, the new scheme should

- ❑ eliminate freely propagating "compressible" short-wave modes that it cannot represent accurately due to spatio-temporal under-resolution,
- ❑ represent with second-order accuracy the "slaved" dynamics of short-wave solution components induced by slow forcing or arising in the form of high-order corrections to long-wave modes, and
- ❑ minimize dispersion for resolved modes.

The first and the last point address the specific dispersion behavior of common second-order implicit time discretizations. While the decision is certainly subjective and depends on the application, at some point the slow down of modes with wave numbers smaller than a certain value is unacceptable. These modes should be eliminated over time in a consistent way. On the other hand, long-wave modes, whose oscillation is well resolved at a fixed position, should be well approximated. The second point refers to the balanced flow on the small scale of the regime, which depends on local source terms and the coupling to the large scale dynamics.

To achieve these goals, a multilevel approach is investigated, which enables the association of different solution components with certain spatial scales. This scale-dependent mode selection mechanism is based on geometric multigrid ideas. Furthermore, selective to each scale, a proper discretization is applied. This results in a robust representation of balanced, slowly forced fast modes on the one hand, and a proper advection of long wave gravity waves on the other hand.

1.1 Purpose and objectives

The main objective of this thesis is to develop a second-order accurate semi-implicit method for low Froude number shallow water flows. It should be a *multiscale* or an *asymptotically adaptive* numerical scheme in that it accurately simulates the regime of long-wave gravity waves traveling over short-range topography and switches into a projection method in the limit of a vanishing Froude number. To achieve this goal, the following questions are addressed in the course of this work:

❏ What are the specific properties of the considered asymptotic regime in the low Froude number limit to be reproduced by the numerical scheme?

❏ What are the discrete-dispersion relations of popular time discretization methods, and how can these be employed in the development of a scale-selective mulitlevel method for the linearized shallow water equations?

❏ Based on an existing projection method for the zero Froude number shallow water equations, how can this method be generalized to a semi-implicit method for low but non-zero Froude number flows?

❏ How can the idea of the multilevel method for linearized shallow water flows be combined with a semi-implicit scheme for the fully nonlinear equations?

To answer these questions, this work begins with the derivation and analysis of the shallow water equations with time-dependent bottom topography in Chapter 2. This includes a discussion of the low Froude number limit and the regime of long-wave gravity waves traveling over short-range topography. Two classical and one newly developed time discretization methods for the linearized shallow water equations are analyzed in Chapter 3. This leads to the construction of a multilevel method for these equations, which incorporates ideas from multigrid techniques. In Chapter 4, a projection method for zero Froude number flows is extended to deal with time dependent bottom topography. Building on the results of the previous chapters, this method is generalized to a semi-implicit method for low Froude number flows and combined with the multilevel method for the linearized equations in Chapter 5. In order to test the numerical properties of the respective methods, numerical tests are are presented in Chapters 3 to 5. In the final part, open questions are discussed and possible approaches for their solution are outlined.

2 Governing equations

In this chapter, the equations used throughout this thesis are presented and discussed. First, the shallow water equations with time dependent bottom topography are derived from the incompressible Euler equations. The assumption of an unsteady bottom topography, which is slightly unusual, is considered to model a source term which is acting on the local flow divergence as outlined in the introduction. Furthermore, the governing equations are linearized, and both the linear and nonlinear versions are analyzed in the limit of a small Froude number. Particularly, the asymptotic limit regime for long-wave shallow water waves passing over short-range topography as presented in Bresch *et al.* (2011) is discussed for topography changing in time. Since the shallow water equations in the limit of the Froude number going to infinity are used in the numerical schemes derived in Chapters 4 and 5, they are also given, and their mathematical structure is discussed. In the final section of this chapter the different approaches for computing low Mach number flows in science and engineering are reviewed and related to the present one.

2.1 Shallow water flows with time dependent bottom topography

The derivation of the shallow water equations can be found in numerous text books (e.g., Pedlosky, 1987; Vallis, 2006). Here the derivation is repeated for the case of non-stationary bottom topography. It will be shown that this addition does not change the structure of the equations. However, in the numerical implementation it must properly be considered as discussed in subsequent chapters.

The main assumptions in the derivation of the inviscid shallow water equations from the Euler equations are the incompressibility and homogeneity of the fluid and the
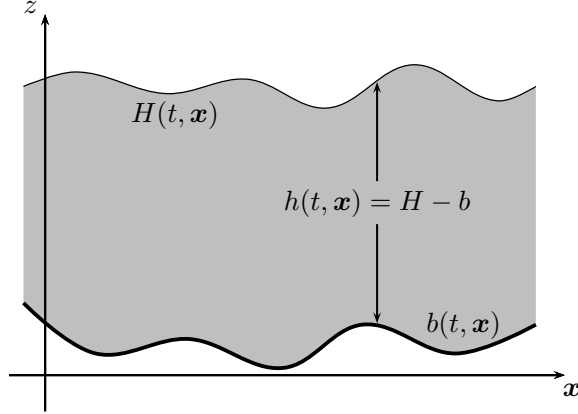
Figure 2.1: The shallow water model with time dependent bottom topography.

assumption of hydrostatic equilibrium, the latter being well satisfied in the presence of a small aspect ratio. Furthermore, at the bottom of the fluid a non-penetration condition is imposed, while at the top a free surface is assumed. Figure 2.1 shows a sketch of the shallow water model in one space dimension. In this thesis no rotational effects are being considered. Thus, they are also neglected in the derivation of the governing equations.

The equations are always given in non-dimensional form. In the dimensional analysis it is assumed that the reference height is independent of the pressure scale height. In fact, the Froude number is set to be of order one in the (incompressible) limit of the Mach number going to zero, which means that the reference height is small compared to the pressure scale height. Thus, the incompressible Euler equations with constant density (normalized to 1 for simplicity) are given by

$$\mathsf{Sr}\,\boldsymbol{v}_t + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} + \nabla p = -\frac{1}{\mathsf{Fr}^2}\boldsymbol{k}$$
$$\nabla\cdot\boldsymbol{v} = 0\;. \tag{2.1}$$

In these equations $\boldsymbol{v} = (u, v, w)$ denotes the velocity, p the pressure and $\boldsymbol{k}$ the unit upward vector in the vertical direction. Since the flow is incompressible, p acts as a Lagrange multiplier to get compliance of the velocity field with the divergence

constraint in (2.1). Furthermore, the dimensionless characteristic numbers

$$\mathsf{Sr} := \frac{\ell_{\mathrm{ref}}}{t_{\mathrm{ref}}\, v_{\mathrm{ref}}} \quad \text{and} \quad \mathsf{Fr} := \frac{v_{\mathrm{ref}}}{\sqrt{g\, h_{\mathrm{ref}}}}$$

are introduced, which are known as the *Strouhal* and the *Froude* number, respectively. While the Strouhal number describes the ratio between the advection time scale $\ell_{\mathrm{ref}}/v_{\mathrm{ref}}$ and the time scale t_{ref} characteristic to the problem, the Froude number defines the ratio between the characteristic flow velocity v_{ref} and the gravity wave speed $\sqrt{g\, h_{\mathrm{ref}}}$.

With the hydrostatic assumption – i.e. gravitational forces are in balance with the pressure gradient, and the fluid does not undergo vertical accelerations – the vertical component of the velocity equation $(2.1)_1$ reduces to

$$\frac{\partial p}{\partial z} = -\frac{1}{\mathsf{Fr}^2} \ ,$$

and the pressure can be written as

$$p(t,x,y,z) = \frac{1}{\mathsf{Fr}^2}(H(t,x,y) - z) + p_a(t,x,y) \ .$$

Here, $H(t,x,y)$ is the top of the fluid at time t and position (x,y) and $p_a(t,x,y)$ is the pressure at this location. Assuming that the weight of the overlying fluid is negligible, the surface pressure can be set to $p_a(t,x,y) \equiv 0$. Therefore, the horizontal pressure gradient is given by

$$\nabla_{\shortparallel} p = \frac{1}{\mathsf{Fr}^2}\nabla_{\shortparallel} H \ ,$$

which is independent of the vertical coordinate. In this equation, the horizontal gradient operator $\nabla_{\shortparallel} := (\partial_x, \partial_y)^T$ was introduced. With this, the horizontal component of the velocity equation can be rewritten as

$$\frac{D\boldsymbol{u}}{Dt} = -\frac{1}{\mathsf{Fr}^2}\nabla_{\shortparallel} H \ , \tag{2.2}$$

where $\boldsymbol{u} = (u, v)$ is the horizontal part of the velocity and

$$\frac{D}{Dt} = \mathsf{Sr}\,\partial_t + (\boldsymbol{v} \cdot \nabla)$$

is the material derivative along particle paths. Since the right hand side of (2.2) does not depend on z, a flow, which is initially independent of z, must stay so. Assuming such initial conditions, the horizontal velocity components u and v are only functions of t, x and y, and the final horizontal velocity equation is given by

$$\mathsf{Sr}\,\boldsymbol{u}_t + (\boldsymbol{u} \cdot \nabla_{\!\shortparallel})\boldsymbol{u} + \frac{1}{\mathsf{Fr}^2}\nabla_{\!\shortparallel}H = \boldsymbol{0}\ . \tag{2.3}$$

Note, that the independence of z in the horizontal velocity components results from the hydrostatic approximation. One could also start by assuming this z-independence and use it as a definition of shallow water flow (Vallis, 2006). However, in realistic applications this assumption would never be fully satisfied.

For the derivation of the height equation, the divergence constraint on the velocity has to be integrated in z, and the boundary conditions at the top and bottom of the fluid are considered. The non-penetration boundary condition states that the relative normal velocity should be zero at the boundary. Thus, let $F_{\mathrm{B}}(t, x, y, z) \coloneqq z - z_{\mathrm{B}}(t, x, y)$, where z_{B} is the position of either the top or the bottom boundary. Then, the boundary is at $F_{\mathrm{B}} = 0$, and the relative velocity at these points must be perpendicular to the gradient of F_{B}, i.e.

$$(u, v, w - \mathsf{Sr}\,\frac{\partial z_{\mathrm{B}}}{\partial t})^T \cdot \nabla F_{\mathrm{B}} = 0\ .$$

For the bottom $b = b(t, x, y)$ and at the top $H = H(t, x, y)$ of the fluid this results in the equations

$$\mathsf{Sr}\,\frac{\partial b}{\partial t} + u\frac{\partial b}{\partial x} + v\frac{\partial b}{\partial y} - w = 0 \tag{2.4}$$

and

$$\mathsf{Sr}\,\frac{\partial H}{\partial t} + u\frac{\partial H}{\partial x} + v\frac{\partial H}{\partial y} - w = 0\ , \tag{2.5}$$

respectively. For the vertical integration of the divergence constraint the general

integration rule

$$\frac{d}{dx} \int_{a(x)}^{b(x)} \varphi(x, z)\, \mathrm{d}z = \int_{a(x)}^{b(x)} \frac{\partial \varphi}{\partial x}(x, z)\, \mathrm{d}z + [\varphi(x, b(x))b'(x) - \varphi(x, a(x))a'(x)]$$

is needed. From this it follows that

$$\int_{b(x,y)}^{H(x,y)} \nabla \cdot \boldsymbol{v}\, \mathrm{d}z = \int_{b(x,y)}^{H(x,y)} (u_x + v_y + w_z)\, \mathrm{d}z$$

$$= \frac{\partial}{\partial x} \int_{b}^{H} u\, \mathrm{d}z + \frac{\partial}{\partial y} \int_{b}^{H} v\, \mathrm{d}z - \left(u \frac{\partial z}{\partial x} + v \frac{\partial z}{\partial y} - w \right)_{z=b}^{H}$$

$$= (hu)_x + (hv)_y - \left(u \frac{\partial z}{\partial x} + v \frac{\partial z}{\partial y} - w \right)_{z=b}^{H} = 0\ ,$$

where $h = H - b$ is the thickness of the fluid. Using the relations (2.4) and (2.5) obtained from the boundary conditions, this yields

$$\mathsf{Sr}\, \frac{\partial H}{\partial t} + (hu)_x + (hv)_y = \mathsf{Sr}\, \frac{\partial b}{\partial t}\ . \tag{2.6}$$

This equation states the conservation of mass (or height). Combining the velocity equation (2.3) with (2.6), a system of conservation laws for height h and momentum $(h\boldsymbol{u})$ is derived, where the momentum equation has a source term due to bottom topography. These are the

Shallow water equations in conservation form:

$$\mathsf{Sr}\, h_t + \nabla \cdot (h\boldsymbol{u}) = 0$$
$$\mathsf{Sr}\, (h\boldsymbol{u})_t + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u}) + \frac{1}{2\mathsf{Fr}^2} \nabla(h^2) = -\frac{1}{\mathsf{Fr}^2} h \nabla b \tag{2.7}$$

In these equations the spatial differential operators are always acting only in the horizontal direction, so the parallel sign is omitted from now on. The "$\circ$" denotes the dyadic product of two vectors. Note that the time dependency of the bottom topography does not result in an additional source term. However, a change in bottom topography either changes the total height H or introduces divergence in the momentum field, as can be seen from (2.6). Furthermore, a change in the gradient of

b directly enters the source term of the momentum equation, leading to a potential disruption of the hydrostatic equilibrium, if the flow was in equilibrium previously.

Remark 2.1 *The shallow water equations without bottom topography have the same mathematical structure as the two-dimensional Euler equations of compressible isentropic gas dynamics, i.e., flows with constant entropy* (Majda, 2003)*. In fact, the latter are given by*

$$\mathsf{Sr}\,\varrho_t + \nabla \cdot (\varrho \boldsymbol{u}) = 0$$

$$\mathsf{Sr}\,(\varrho \boldsymbol{u})_t + \nabla \cdot \left(\varrho \boldsymbol{u} \circ \boldsymbol{u} + \frac{\hat{\kappa}}{\mathsf{M}^2}\,\varrho^\gamma\,\boldsymbol{I} \right) = 0$$

with constants $\hat{\kappa}$ and $\gamma > 1$, and $\boldsymbol{I}$ being the 2×2 identity matrix. The dimensionless characteristic quantity $\mathsf{M} := v_{\mathrm{ref}}/\sqrt{p_{\mathrm{ref}}/\rho_{\mathrm{ref}}}$ is the Mach number. By replacing the density ϱ with h, $\hat{\kappa}/\mathsf{M}^2$ with $1/(2\mathsf{Fr}^2)$, and setting $\gamma = 2$, the shallow water equations without source terms are recovered. ◁

Although the shallow water equations are mathematically equivalent to a special case of the compressible Euler equations, an incompressible fluid is modeled in the former case. However, as we also saw in the above remark, the Froude number in the shallow water equations takes the role of the Mach number in the Euler equations, the latter being a measure of the compressibility of the fluid. Therefore, some "compressibility" effects can be also modeled by the shallow water equations, where the importance of the "compressibility" depends on the associated scales of the fluid motion. In large scale atmospheric applications, a typical flow velocity is 10 m/s and the depth of the atmosphere is given by the pressure scale height, which is approximately 10 km. This results in a Froude number $\mathsf{Fr} \approx 0.03 \ll 1$, and the "compressibility" effects associated with the nonlinear nature of gravity waves will play a minor role in this regime. Furthermore, in practical applications the phenomena which are associated with the advective time scale of the fluid are of major interest. Choosing this scale as reference, this leads to a Strouhal number of order one. This quantity is set to $\mathsf{Sr} = 1$ from now on.

It is worth mentioning that for smooth solutions the shallow water equations satisfy some further conservation principles. First, the so-called *potential vorticity* is conserved along particle paths: with the definition of the vorticity ω in two space

dimensions by

$$\omega := \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$$

it can be shown that

$$\frac{D}{Dt}\left(\frac{\omega}{h}\right) = 0 \, ,$$

where $q := \omega/h$ is the potential vorticity (see, e.g., Majda, 2003). The consequence is that height and vorticity are proportional along particle trajectories.

Furthermore, the shallow water equations imply the conservation of energy, which is a principle of great physical and mathematical importance. Let us define the kinetic and potential energy by

$$E_{\mathrm{kin}} := \frac{h\boldsymbol{u} \cdot \boldsymbol{u}}{2} \quad \text{and} \quad E_{\mathrm{pot}} := \frac{h}{\mathsf{Fr}^2}\left(\frac{h}{2} + b\right) \, .$$

Then, the total energy $E := E_{\mathrm{kin}} + E_{\mathrm{pot}}$ satisfies

$$\frac{\partial E}{\partial t} + \nabla_{\shortparallel} \cdot \left(\left(E + \frac{h^2}{2\mathsf{Fr}^2}\right)\boldsymbol{u}\right) = \frac{h}{\mathsf{Fr}^2}\frac{\partial b}{\partial t} \, .$$

Thus, for a bounded domain and time independent topography energy is conserved and can only change by some flux over the boundary of the domain. However, a change in mean bottom topography also implies a change of the total energy.

2.2 Linearized long-wave shallow water flow

The numerical scheme, which is constructed in Chapter 5, is based on the idea that the nonlinear slow components and the fast linear components of the flow can be treated to some extent separately. In particular, a new multilevel scheme will be proposed for the linear part in Chapter 3. In this section, the one-dimensional version of the shallow water equations is linearized around a state at rest (H_0, u_0) with $u_0 = 0$, and the resulting system is analyzed for the asymptotic regime of long-wave shallow water waves passing over short-range topography. This is the same regime as it was considered in Bresch *et al.* (2011) for the fully nonlinear shallow water equations, the

latter being revisited in Section 2.4. Thus the asymptotic analysis is very similar, except that here all the nonlinear terms can be omitted, but the time dependency of the bottom topography has to be taken into account.

For the linearization, let us set

$$
\begin{aligned}
h(t, x) &= H_0 + \varepsilon h'(t, x) - b(t, x) \\
b(t, x) &= \varepsilon b'(t, x) \\
u(t, x) &= \varepsilon u'(t, x) \ ,
\end{aligned}
$$

where H_0 is the constant background height, and the quantities with primes are the perturbations around the linearized state. Omitting all terms of order ε^2, this results in the

Linearized shallow water equations:

$$
\begin{aligned}
h'_t + H_0 u'_x &= b'_t \\
u'_t + \frac{1}{\mathsf{Fr}^2} h'_x &= 0 \ .
\end{aligned}
\tag{2.8}
$$

By looking for traveling wave solutions, which are of the form $(h', u')(t, x) = \mathrm{Re}[(h'_0, u'_0) \exp(i(\omega t - \kappa x))]$, the dispersion relation

$$
\omega^2 - \frac{H_0}{\mathsf{Fr}^2} \kappa^2 = 0
\tag{2.9}
$$

is obtained. Thus, $\omega(\kappa) = \pm c\kappa$ where $c \coloneqq \frac{\sqrt{H_0}}{\mathsf{Fr}}$, so that in the continuous system (2.8) all waves travel with the same velocity, $\pm c$, without dispersion. These waves are known as *(external) gravity waves*, due to their origin from the displacement of the fluid height. From now on the primes of the perturbation terms are omitted in this section. As in the nonlinear case, the linearized shallow water equations imply conservation of energy. From (2.8) it is easy to derive the balance equation

$$
\frac{\partial}{\partial t} \left(\frac{h^2}{2\,\mathsf{Fr}^2} + H_0\,\frac{u^2}{2} \right) + \frac{H_0}{\mathsf{Fr}^2} \frac{\partial}{\partial x} (hu) = \frac{h}{\mathsf{Fr}^2} b_t \ ,
\tag{2.10}
$$

and it follows conservation of the global pseudo-energy

$$E_{\text{global}}(t) = \int_\Omega \frac{h^2}{2\,\mathsf{Fr}^2} + H_0\,\frac{u^2}{2}\,\mathrm{d}x \tag{2.11}$$

up to flow across the boundary of the domain Ω and variations of the mean bottom topography in time.

Remark 2.2 *The equations corresponding to the linearized shallow water equations in the field of gas dynamics are the*

Equations for linear acoustics:

$$\begin{aligned} m_t + p_x \quad &= 0 \\ p_t + a^2 m_x &= 0\ . \end{aligned}$$

Here, $p = p(t,x)$ and $m = m(t,x)$ are the pressure and momentum perturbations around a state at rest, respectively, and a is the speed of sound. ◁

In the following, a multiple-scales asymptotic analysis for the linearized shallow water equations in the limit of a small Froude number $\mathsf{Fr} \ll 1$ is conducted (for details about this technique see, e.g., Schneider (1978); Klein *et al.* (2011)). As mentioned above, we are interested in long-wave shallow water waves traveling over rough topography. Therefore, a second large space scale $\xi = \mathsf{Fr}\,x$ is introduced, and the bottom topography $b = b(t,x,\xi)$ can depend on time and both space scales. Height and velocity are assumed to have the multiple-scales expansion

$$(h,u)(t,x;\mathsf{Fr}) = \sum_{i=0}^N \mathsf{Fr}^i (h,u)^{(i)}(t,x,\xi) + o\!\left(\mathsf{Fr}^N\right)\ .$$

Note, that by using this ansatz, each spatial derivative of an asymptotic function $\varphi^{(i)}$ translates into

$$\left.\frac{\partial \varphi^{(i)}}{\partial x}\right|_{\mathsf{Fr}} = \frac{\partial \varphi^{(i)}}{\partial x} + \mathsf{Fr}\frac{\partial \varphi^{(i)}}{\partial \xi}$$

for fixed Froude number Fr. Inserting the above ansatz into the equations (2.8) results

in

$$h_t^{(0)} + \mathsf{Fr}\, h_t^{(1)} + \cdots + H_0(u_x^{(0)} + \mathsf{Fr}\, u_\xi^{(0)} + \mathsf{Fr}\, u_x^{(1)} + \dots) = b_t\ ,$$

$$u_t^{(0)} + \mathsf{Fr}\, u_t^{(1)} + \cdots + \frac{1}{\mathsf{Fr}^2}(h_x^{(0)} + \mathsf{Fr}\, h_\xi^{(0)} + \mathsf{Fr}\, h_x^{(1)} +$$

$$\mathsf{Fr}^2\, h_\xi^{(1)} + \mathsf{Fr}^2\, h_x^{(2)} + \dots) = 0\ .$$

The leading order dynamics are then determined by collecting terms of equal order in Fr, and setting each sum separately to zero. This leads to the following system of equations:

$$\mathcal{O}\!\left(\mathsf{Fr}^{-2}\right): \qquad\qquad\qquad h_x^{(0)} = 0$$

$$\mathcal{O}\!\left(\mathsf{Fr}^{-1}\right): \qquad\qquad h_\xi^{(0)} + h_x^{(1)} = 0$$

$$\mathcal{O}\!\left(\mathsf{Fr}^{0}\right): \qquad\qquad h_t^{(0)} + H_0 u_x^{(0)} = b_t$$

$$u_t^{(0)} + h_\xi^{(1)} + h_x^{(2)} = 0$$

$$\mathcal{O}\!\left(\mathsf{Fr}^{1}\right): \qquad h_t^{(1)} + H_0(u_\xi^{(0)} + u_x^{(1)}) = 0$$

The leading order equation states that $h^{(0)}$ is independent of x, and, averaging the next order equation over an interval $[-r, r]$ in x, it follows that

$$h_\xi^{(0)} + \frac{1}{2r}\left(h^{(1)}(t, r, \xi) - h^{(1)}(t, -r, \xi)\right) = 0\ .$$

To obtain successively smaller correction terms in the asymptotic expansion, the *sub-linear growth condition* has to be enforced on $h^{(1)}$, which is $\mathsf{Fr}\, h^{(1)}(t, \xi/\mathsf{Fr}, \xi) = o(1)$ for fixed ξ and $\mathsf{Fr} \to 0$ (Schneider, 1978; Klein *et al.*, 2011). This implies $h^{(1)}(t, x, \xi) = o(|x|)$ for $|x| \to \infty$, and therefore

$$h_\xi^{(0)} = \mathcal{O}(1/r) \cdot o(r) = o(1) \quad \text{for } r \to \infty\ ,$$

which means that $h^{(0)}$ is only a function of time, i.e. $h^{(0)} = h_0(t)$. Reapplied to the equation of order Fr^{-1}, it follows that also $h^{(1)} = h_1(t, \xi)$ is independent of x. The same procedure (averaging and enforcement of sub-linear growth) is adopted to the

subsequent equations. Averaging the height equation of order Fr^0 results in

$$\frac{dh_0}{dt} = \bar{b}_t \; ,$$

where the line over a quantity denotes the average in the x-coordinate. Subtracting this from the original equation yields

$$H_0 u_x^{(0)} = H_0 \widetilde{u^{(0)}}_x = b_t - \bar{b}_t = \widetilde{b}_t$$

with $\widetilde{u^{(0)}} = u^{(0)} - \overline{u^{(0)}}$. The velocity equation of the same order decomposes into

$$\overline{u^{(0)}}_t + h_\xi^{(1)} = 0 \qquad \text{and} \qquad \widetilde{u^{(0)}}_t + h_x^{(2)} = 0 \; .$$

Finally, the assumption of sub-linear growth applied to the height equation of order Fr^1 leads to

$$h_t^{(1)} + H_0 \overline{u^{(0)}}_\xi = 0 \; .$$

Summarizing the results from this analysis, two sets of equations are obtained, one for the large scale ξ and another one for the small scale x. These are the

Long-wave equations:

$$\begin{aligned} h_t^{(1)} + H_0 \overline{u^{(0)}}_\xi &= 0 \; , \\ \overline{u^{(0)}}_t + h_\xi^{(1)} &= 0 \; , \end{aligned} \tag{2.12}$$

and the

Balanced small scale flow:

$$\begin{aligned} \widetilde{u^{(0)}}_t + h_x^{(2)} &= 0 \; , \\ H_0 u_x^{(0)} = H_0 \widetilde{u^{(0)}}_x &= \widetilde{b}_t \; . \end{aligned} \tag{2.13}$$

On the large scale, essentially the linearized shallow water equations are recovered. On this scale the gravity wave speed $c_\xi = \sqrt{H_0}$ is of order 1. On the contrary, the small scale flow is governed by a divergence constraint on the velocity field. The

height perturbation $h^{(2)}$ takes the role of a Lagrange multiplier. Note, that in the linear case the systems are fully decoupled, and there is no interaction between the scales. Differentiating the first equation of the small scale system with respect to x and the second with respect to t, the height perturbation $h^{(2)}$ is given by the elliptic equation

$$-h_{xx}^{(2)} = \widetilde{u^{(0)}}_{tx} = \frac{1}{H_0}\widetilde{b}_{tt} \ . \tag{2.14}$$

The asymptotic scaling for the velocity in this regime is given by $u \sim 1$ as $\mathsf{Fr} \to 0$. For the height we have $h - h_0(t) \sim \mathsf{Fr}$ on the large scale and $h - h_0(t) \sim \mathsf{Fr}^2$ on the small scale, respectively. This scaling should be reproduced by a numerical scheme, especially when $\Delta t \gg \frac{\Delta \xi}{\sqrt{H_0}} = \frac{\mathsf{Fr}\Delta x}{\sqrt{H_0}}$, the latter corresponding to large Courant numbers with respect to gravity waves for the time integration in the present model problem. Furthermore, solutions that are only triggered by the source term should have the structure given by equation (2.14) up to small perturbations. If this scaling is applied to the governing equations (2.8) (i.e., $u = \tilde{u}$ and $h = \mathsf{Fr}^2\tilde{h}$), the resulting system is

$$\begin{aligned} \mathsf{Fr}^2\tilde{h}_t + H_0\tilde{u}_x &= b_t(t, x, \xi) \ , \\ \tilde{u}_t + \tilde{h}_x \ \ &= 0 \ , \end{aligned} \tag{2.15}$$

where the first term in the first equation becomes singular for $\mathsf{Fr} \to 0$.

2.3 The "lake equations"

Let us now turn back to the fully nonlinear shallow water equations (2.7) with bottom topography varying in time and analyze their asymptotic behavior in the limit $\mathsf{Fr} \to 0$. First, the single scale limit is considered, in which bottom topography $b = b(t, \boldsymbol{x})$, height and velocity

$$(h, \boldsymbol{u})(t, x; \mathsf{Fr}) = \sum_{i=0}^{N} \mathsf{Fr}^i(h, \boldsymbol{u})^{(i)}(t, \boldsymbol{x}) + \mathcal{O}\left(\mathsf{Fr}^N\right) \ .$$

only depend on one space and time scale, respectively. Without time dependent bottom topography, this regime has been analyzed in Bresch *et al.* (2011) and without bottom topography in Vater (2005). Therefore, only the results and the differences

compared to the case with stationary topography are presented. The leading order system is given by

$$\nabla \cdot (h\boldsymbol{u})^{(0)} = -\frac{dH^{(0)}}{dt} + \frac{\partial b}{\partial t} \; , \qquad (2.16)$$
$$(h\boldsymbol{u})_t^{(0)} + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{(0)} + h^{(0)}\nabla h^{(2)} = 0 \; ,$$

where the total height is defined by

$$H^{(0)}(t) \coloneqq h^{(0)}(t, \boldsymbol{x}) + b(t, \boldsymbol{x}) \; .$$

Compared to Bresch *et al.* (2011) an additional forcing term b_t appears in the continuity equation $(2.16)_1$, which acts on the local divergence of the flow. The evolution of $H^{(0)}$ in time is obtained by integrating the continuity equation over the entire domain Ω, which leads to

$$\frac{dH^{(0)}}{dt} = -\frac{1}{|\Omega|} \int_{\partial\Omega} (h\boldsymbol{u})^{(0)} \cdot \boldsymbol{n} \, \mathrm{d}\sigma + \frac{1}{|\Omega|} \int_{\Omega} \frac{\partial b}{\partial t} \, \mathrm{d}\boldsymbol{x} \; , \qquad (2.17)$$

i.e., the change in total fluid height is given by the integrated height flux across the boundary of the domain and the mean evolution of the bottom topography. Note, that due to the local time dependence of b also $h^{(0)}$ is changing locally in time, and not only globally by $H_t^{(0)}$.

The equations (2.16) are known as the *lake equations* and have already essentially been described by Greenspan (1968). They model a balanced flow without gravity waves. The second-order height $h^{(2)}$ acts again as a Lagrange multiplier in order to have the velocity field in compliance with the divergence constraint. Levermore *et al.* (1996) proved the global well-posedness of these equations, i.e. existence and uniqueness of solutions and their continuous dependence on the data. Furthermore, Oliver (1997) proved that the lake equations are a good approximation to the *rigid-lid equations* (the three-dimensional Euler equations with a rigid-lid upper boundary condition) in a horizontally periodic basin with bottom topography for a finite interval of time.

2.4 Long-wave gravity waves passing over short-range topography

In addition to the single scale limit of the shallow water equations, the regime of long-wave shallow water waves passing over topography is of interest, which was already considered for the linear case in Section 2.2. Once again, a second large space scale $\boldsymbol{\xi} = \mathsf{Fr}\,\boldsymbol{x}$ is introduced, which resolves the distance a gravity wave traverses on the considered time scale. For the bottom topography $b = b(t, \boldsymbol{x}, \boldsymbol{\xi})$ we allow for variations on both space scales, and height and velocity are expressed in the multiple-scales expansion

$$(h, \boldsymbol{u})(t, x; \mathsf{Fr}) = \sum_{i=0}^{N} \mathsf{Fr}^i (h, \boldsymbol{u})^{(i)}(t, \boldsymbol{x}, \boldsymbol{\xi}) + \mathcal{O}\left(\mathsf{Fr}^N\right) \ .$$

As stated above, this regime has been also discussed in Bresch *et al.* (2011), but without time dependent bottom topography. So again, only the results are given, and differences are pointed out where they occur. The leading order system is separated into two subsystems representing the long-wave and the short-wave components of the flow. They are given by the

Long-wave equations for rough topography:

$$\overline{(h\boldsymbol{u})^{(0)}}_t + \overline{h^{(0)}} \nabla_{\boldsymbol{\xi}} h^{(1)} = \overline{h^{(2)} \nabla_{\boldsymbol{x}} h^{(0)}} \ ,$$
$$h_t^{(1)} + \nabla_{\boldsymbol{\xi}} \cdot \overline{(h\boldsymbol{u})^{(0)}} = 0 \ ,$$

(2.18)

and the

Balanced small scale flow for rough topography:

$$\widetilde{(h\boldsymbol{u})^{(0)}}_t + \nabla_{\boldsymbol{x}} \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{(0)} + h^{(0)} \widetilde{\nabla_{\boldsymbol{x}} h^{(2)}} = -\widetilde{h^{(0)}} \nabla_{\boldsymbol{\xi}} h^{(1)} \ ,$$
$$\nabla_{\boldsymbol{x}} \cdot (h\boldsymbol{u})^{(0)} = \nabla_{\boldsymbol{x}} \cdot \widetilde{(h\boldsymbol{u})^{(0)}} = \widetilde{b}_t \ .$$

(2.19)

The leading order height is given by

$$h^{(0)}(t, \boldsymbol{x}, \boldsymbol{\xi}) = H^{(0)}(t) - b(t, \boldsymbol{x}, \boldsymbol{\xi}) \ ,$$

where $H^{(0)}$ is the total surface height of the fluid and $\frac{dH^{(0)}}{dt} = \bar{b}_t$. The next order of the height $h^{(1)} = h_1(t, \boldsymbol{\xi})$ is independent of x.

Compared to the linear case (i.e. equations (2.12) and (2.13)) the two systems (2.18) and (2.19) are coupled. The large-scale flow is given by the linearized shallow water equations, which involve non-balanced free surface waves. It is driven by a source term arising from the small scale flow in the momentum equation. This source represents the accumulated pressure force, which results from the small-scale flow across the rough topography. In the opposite direction, large scale height gradients acting on the rough topography induce small scale momentum. This modifies the otherwise balanced small scale flow.

The difference to Bresch *et al.* (2011) when considering non-stationary bottom topography is that a source term acting on the local divergence of the flow arises again. It is generated by local variations of the bottom topography in time. Furthermore, the changes of the mean in b over time induce a change in the total surface height $H^{(0)}$, and the signal speed of the long-wave gravity waves is changing not only in space, but also in time.

Similar asymptotic regimes were studied in Klein (1995) concerning weakly compressible flows with small-scale entropy and vorticity, in Le Maître *et al.* (2001) for modeling ocean flows, and in the context of atmospheric circulation near the equator in Majda and Klein (2003).

Remark 2.3 *Some more insight into this regime can be gained, when we restrict ourselves to one-dimensional flows, although it is not clear to which degree these findings are also valid for higher dimensions. In one space dimension the momentum equation of order* Fr^0 *is given by*

$$u_t^{(0)} + u^{(0)} u_x^{(0)} + h_x^{(2)} + h_\xi^{(1)} = 0 \ .$$

The sub-linear growth condition then implies that

$$\overline{u^{(0)}}_t = -h^{(1)}_\xi \qquad and \qquad \widetilde{u^{(0)}}_t + u^{(0)} u^{(0)}_x + h^{(2)}_x = 0 \ .$$

Differentiating the large scale equation with respect to t and equation $(2.18)_2$ *with respect to* ξ *yields*

$$\overline{u^{(0)}}_{tt} = -h^{(1)}_{t\xi} = \overline{(hu)^{(0)}}_{\xi\xi} \ . \tag{2.20}$$

On the other hand, in one dimension we can integrate the divergence constraint $(2.19)_2$ *in x to get*

$$(hu)^{(0)} = \widetilde{B}_t + f(t,\xi) \quad with \quad \widetilde{B} := \int_0^x \widetilde{b} \, \mathrm{d}x' \ ,$$

where the function f represents the large scale part of $(hu)^{(0)}$. *Therefore, the momentum splits into*

$$\overline{(hu)^{(0)}} = f(t,\xi) \quad and \quad \widetilde{(hu)^{(0)}} = \widetilde{B}_t \ .$$

The leading order velocity can be obtained by dividing the momentum by the (positive) leading order height, i.e.

$$u^{(0)} = \frac{(hu)^{(0)}}{h^{(0)}} = \frac{\widetilde{B}_t + f(t,\xi)}{H^{(0)} - b} \ ,$$

and its large scale component is

$$\overline{u^{(0)}} = \overline{\left(\frac{\widetilde{B}_t}{H^{(0)} - b}\right)} + \overline{\left(\frac{1}{H^{(0)} - b}\right)} f(t,\xi) =: \overline{C} + \overline{D} f(t,\xi) \ .$$

If the representations of $\overline{u^{(0)}}$ *and* $\overline{(hu)^{(0)}}$ *are inserted into* (2.20), *a modulated wave equation is obtained for the function f. It is given by*

$$(\overline{D} f_{tt} - f_{\xi\xi}) + 2\overline{D}_t f_t + \overline{D}_{tt} f = -\overline{C}_{tt} \ ,$$

which can even be solved numerically in order to get an asymptotic reference solution for comparison with numerical simulations of the full equations. ◁

2.5 The "pressureless equations"

In this section a limit regime of the shallow water equations is discussed, which is completely different to the considerations made so far. In particular, the limit $\mathsf{Fr} \to \infty$ is studied. The resulting equations are usually used to model so-called sticky particles, which explain the formation of large scale structures in the universe. However, in this work they will be employed in the construction of the numerical scheme for the approximate solution of zero and low Froude number shallow water flows (see Chapters 4 and 5 for details).

In the limit of $\mathsf{Fr} \to \infty$ the pressure gradient in the momentum equation simply vanishes and the limit equations are

$$
\begin{aligned}
h_t + \nabla \cdot (h\boldsymbol{u}) &= 0 \ , \\
(h\boldsymbol{u})_t + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u}) &= \boldsymbol{0} \ .
\end{aligned}
\tag{2.21}
$$

This system is equivalent to the so-called *equations of pressureless gas dynamics* and expresses the conservation of mass and momentum in the absence of pressure forces. It has been extensively studied in the literature (see e.g. Bouchut, 1994; Bouchut *et al.*, 2003; LeVeque, 2004, and references therein).

To analyze the characteristic structure of (2.21), let us consider the shallow water equations in one space dimension. These can be written in the quasilinear form

$$
\begin{pmatrix} h \\ hu \end{pmatrix}_t + F' \cdot \begin{pmatrix} h \\ hu \end{pmatrix}_x = 0 \ , \quad \text{where} \quad F' = \begin{pmatrix} 0 & 1 \\ -u^2 + \frac{1}{\mathsf{Fr}^2}h & 2u \end{pmatrix} \ .
\tag{2.22}
$$

For finite Froude numbers the shallow water equations are hyperbolic, and the Jacobian F' of the flux has eigenvalues $\lambda = u \pm \frac{1}{\mathsf{Fr}}\sqrt{h}$. However, as the Froude number goes to infinity, system (2.22) degenerates and has only one eigenvalue $\lambda_{\text{lim}} = u$ with algebraic multiplicity 2 associated to the eigenvector $(1, u)^T$. As a result the limit system features density concentrations known as delta shocks and the occurrence of the vacuum state.

The solution of the Riemann problem for finite Froude number consists of two distinct waves, which can be shocks or rarefaction waves. In the limit of a vanishing pressure, however, the solution consists either of two contact discontinuities with

vacuum in between (when $u_l < u_r$), or of a finite jump combined with δ-singularities in height and momentum if $u_l > u_r$. This δ-shock can be shown to have the propagation speed

$$u_\delta = \frac{\sqrt{h_l}u_l + \sqrt{h_r}u_r}{\sqrt{h_l} + \sqrt{h_r}} \; , \qquad (2.23)$$

which corresponds to the usual Roe-average for the velocity (cf. LeVeque, 2002, p. 321). The exact solution of the Riemann problem can be found, e.g., in Bouchut *et al.* (2003).

2.6 The low Mach number problem in science and engineering

The scheme for the computation of low Froude number shallow water flows, which will be derived in Chapters 3 to 5, is an extension of a projection method for the zero Froude number limit equations. As discussed in Section 2.1, this is equivalent to the low Mach number limit of the Euler equations for compressible gas dynamics. In this last section of the chapter, the different approaches for the solution of the fully compressible equations in meteorological and engineering applications are reviewed and related to the present one.

The singular behavior of the compressible Euler equations in the low Mach number limit was analyzed by several authors. Klainerman and Majda (1981) and Ebin (1982) rigorously proved that the weakly compressible equations for a isentropic fluid converge to the constant-density incompressible equations under suitable assumptions. These results were extended by Schochet (1988) to variable-density incompressible flows (see also the recent review by Schochet, 2005). The singularity is expressed by a change of type of the equations from hyperbolic to mixed hyperbolic-elliptic. Due to the above mentioned equivalence of the Euler equations of isentropic gas dynamics to the shallow water equations these results should also hold for the latter ones.

As described in the introduction, there have been different approaches to solve the equations governing atmospheric flow in meteorological problems. The application of *split-explicit* or *multirate* methods was first proposed by Klemp and Wilhelmson (1978) and Gadd (1978), who used a leapfrog or Lax-Wendroff scheme for the integration of

the slow wave components and a forward-backward scheme for the fast components. Later, this technique was also applied in combination with Runge-Kutta methods for the slow waves (Wicker and Skamarock, 1998, 2002). However, as described in Skamarock and Klemp (1992), these methods have some stability issues, and often a so-called divergence damping must be applied in order to stabilize them. This usually reduces the accuracy of the method to first order. To circumvent this problem Jebens *et al.* (2009) proposed a scheme based on peer methods, which are a generalization of classical time step methods and combine positive features of Runge-Kutta and linear multistep methods. With the resulting three-stage scheme it was possible to stably integrate the compressible equations without any artificial damping. Also notable are multirate infinitesimal step methods (Wensch *et al.*, 2009), which generalize classical split-explicit methods by the inclusion of fixed tendencies from previous stages. This approach is different compared to the present one in that the linear part is treated differently. While the explicit discretization of the fast components might result in a better representation of the short-wave modes compared to an implicit treatment, it is often difficult to obtain a stable approximation of second or higher order.

The second approach in the class of "split" methods is the application of *semi-implicit* or *linearly implicit* schemes. The semi-implicit method was introduced in numerical weather prediction by Robert (1969) and Kwizak and Robert (1971). In these works trapezoidal differencing for the fast components was combined with a leapfrog scheme for the slow components. However, the computational mode present in the leapfrog scheme introduces an instability, which has to be controlled, e.g., by a Robert-Asselin time filtering (Robert, 1966; Asselin, 1972). This degrades the method to first order. Therefore, Giraldo (2005) used the BDF(2) method instead, to obtain second-order accuracy. Also, the semi-implicit method was successfully combined with the semi-Lagrangian method (Robert, 1982; Benoit *et al.*, 1997), which can allow for even larger timesteps. Recently, Durran and Blossey (2012) developed semi-implicit (or *implicit-explicit* (IMEX)) linear multistep methods which are based on Adams and backward differentiation schemes to achieve up to third order accuracy with almost the same CFL numbers as the classical semi-implicit methods.

The application of linear implicit Rosenbrock-type methods has been reported in Knoth (2006) for the solution of atmospheric flow problems at different scales. Also, St-Cyr and Neckels (2009) used such a method for the development of a

discontinuous Galerkin mesoscale model. In Ullrich and Jablonowski (2012), a Rosenbrock method was combined with an explicit Runge-Kutta scheme for the application in nonhydrostatic atmospheric models. In this approach the vertical components were treated by a Rosenbrock method, while the Runge-Kutta scheme was used for the horizontal direction to circumvent the stringent time step restriction caused by the large aspect ratio between horizontal and vertical grid spacing. To circumvent the severe time step restriction induced by cut cells for the representation of orography, Jebens *et al.* (2012) introduced a linearly implicit two-stage peer method, which is of second-order accuracy.

These approaches are very similar to the present one in that they combine an implicit treatment of the fast components of the flow with an explicit treatment of the slow part. However, most if not all of the described methods are based rather on a direct splitting of the equations and not on the extension of a projection method. Also, to the best of the author's knowledge there have been no attempts to correctly approximate meteorologically relevant asymptotic multi-scale regimes such as the regime of long-wave acoustic waves interacting with slow advection.

A completely different approach is the use of fully implicit methods for the solution of the compressible flow equations for meteorological problems. Since their application involves the solution of a nonlinear system, they are often thought to be less efficient than, for example, semi-implicit methods. However, some attempts in this direction can be also found in the literature. Reisner *et al.* (2001, 2003, 2005) used a Jacobian-free Newton-Krylov (JFNK) method, which is a combination of a nonlinear outer Newton-based iteration and a linear inner conjugate residual (Krylov) iteration. This method was combined with a "physics-based" preconditioner, which uses the semi-implicit method to solve the governing equations. A related approach is reported in Evans *et al.* (2010).

Low Mach number flows also arise in a number of engineering applications, such as the simulation of thermo-acoustic systems as they occur in industrial gas turbines or the computation of fire events in car tunnels. The methods discussed in the following mostly originate from this research area. The singular behavior of the Euler equations is probably one of the main reasons why historically different methods were developed for compressible problems including shocks and incompressible flows. Furthermore, the following difficulties were identified in computing weakly compressible flows (cf.

Klein, 1995; Schneider *et al.*, 1999):

- ❏ spatial pressure variations vanish as the Mach number $\mathsf{M} \to 0$, but they do affect the velocity field at leading order (the *dynamic range problem*),

- ❏ the spatial homogeneity of leading order pressure implies an elliptic divergence constraint for the mass flux, and

- ❏ the eigenvalues of the Jacobian flux matrix become singular (the *signal speed problem*).

For explicit shock capturing methods, originally designed for compressible flows, this means that they suffer from a CFL time step restriction (Courant *et al.*, 1928) with $\Delta t \leq \mathcal{O}(\mathsf{M})$. The dynamic range problem can be cured by working on perturbations to a local reference state (Sesterhenn *et al.*, 1999). Another problem that arises for these kinds of methods is that in certain circumstances local truncation errors can grow with vanishing Mach numbers. This problem was recently revisited by Rieper (2008) and a solution for Godunov-type methods was presented in Rieper (2011).

Methods for the computation of incompressible flows are commonly based on the predictor-corrector principle. In a first step a provisional velocity field is computed, which is corrected in a second step subject to the divergence constraint of incompressible flows. Furthermore, they are divided into methods for computing stationary and instationary problems. For stationary problems, the divergence constraint is attained in an iterative process. Popular representatives are the artificial compressibility method (Chorin, 1967) and SIMPLE-type methods (Semi-Implicit Method for Pressure-Linked Equations) introduced by Patankar and Spalding (1972). For unsteady incompressible flows projection methods are usually employed, which were first introduced by Chorin (1968) and Temam (1968). There have been many attempts to apply those methods to weakly compressible flows. Karki and Patankar (1988) modified the SIMPLE algorithm in that they added an additional hyperbolic term into the pressure correction equation (see also Demirdžić *et al.*, 1993). First extensions of the projection method can be found in Casulli and Greenspan (1984) for a staggered grid and in Patnaik *et al.* (1987) for a conservative formulation with collocated variables. More recent developments are given by van der Heul *et al.* (2003) and Nerinckx *et al.* (2005).

In Klein (1995) the author performed a multiple-scales asymptotic analysis for the Euler equations at low Mach numbers. It includes variable density flows and two space scales which are associated to small-scale entropy fluctuations and long-wave acoustic waves. Based on the results of the asymptotic analysis, a flow solver for one-dimensional weakly compressible flows was constructed. This method incorporated multiple pressure variables (MPV) and discrete large scale differencing and averaging procedures. The idea was extended to higher dimensions by Geratz (1997), and it was further applied for the extension of SIMPLE-type and projection methods in Munz *et al.* (2003) and Park (2003). Since these schemes were supposed to correctly approximate the considered asymptotic limit, they were referred to as *asymptotically adaptive numerical methods* (Klein, 2000; Klein *et al.*, 2001). Based on the same asymptotic analysis projection-type methods were developed for the variable-density incompressible Euler equations (Schneider *et al.*, 1999), the zero Froude number shallow water equations (Vater, 2005) and the anelastic and pseudo-incompressible models (Klein, 2009). The present scheme can be viewed as a further development of these schemes and builds on several ideas already mentioned in the cited references.

3 A multilevel method for long-wave linear shallow water flows

In the numerical scheme to be developed for computing low Froude number shallow water flows the fast traveling gravity waves should be integrated by an implicit time discretization, to overcome the severe time step restriction in explicit methods. However, as indicated in the introduction such implicit schemes often result in an undesirable numerical dispersive behavior. In this chapter, this phenomenon is studied in detail for some popular classical discretization schemes used in meteorological applications, and, based on the findings, a new multilevel scheme is constructed to resolve some of the issues.

This study was already carried out for the linear acoustics equations in Vater *et al.* (2011). Here, the findings are recalled for the case of the linearized shallow water equations

$$
\begin{aligned}
h_t + H_0 u_x &= b_t \ , \\
u_t + \frac{1}{\mathsf{Fr}^2} h_x &= 0 \ ,
\end{aligned}
\tag{3.1}
$$

as they were introduced in Section 2.2. Referring to the introduction, the new scheme should eliminate freely propagating short-wave gravity-wave modes that it cannot resolve in time, but minimize the dispersion error for resolved modes. The latter are those long-wave modes, which are advected with an "acceptable" dispersion and amplitude error. At the same time, the resulting balance of short-wave solution components generated by slow forcing should be represented with second-order accuracy. To achieve these goals, the scheme incorporates

❏ a scale-dependent mode selection mechanism based on geometric multigrid ideas and

❑ a scale-selective application of proper discretizations for the robust representation of balanced, slowly forced fast modes.

Multilevel schemes have been used elsewhere before, although they often have been only used for two or three different levels, in practice. Especially, they have been developed for the Navier-Stokes equations in the computation of turbulent flows (see e.g., Dubois *et al.*, 2004). Another example is Dubois *et al.* (2005), where a spectral multilevel method for the computation of the shallow water equations is proposed. This scheme increases time step stability compared to an explicit method, while minimizing the dispersive error introduced by implicit discretizations.

3.1 Implicit second-order staggered grid schemes

Here the point of departure of the present developments is described, classical implicit second-order time integration schemes employing central differences on a staggered grid. It should be noted that there are other popular discretizations such as multistage schemes. However, these methods usually require at least two linear systems to be solved in order to achieve second-order accuracy. Because of this additional cost per time step, the analysis is confined to schemes which only require the solution of one linear system per time step. For simplicity and with a view to the application of the developed method to a semi-implicit discretization of the fully nonlinear shallow water equations in Chapter 5, a staggered grid is chosen, but the key ideas should transfer directly to collocated grid arrangements as well.

3.1.1 Standard time integration schemes

The investigated time integration schemes are the implicit trapezoidal rule and the BDF(2) scheme (backward differentiation formula of second order). These are discretizations commonly used in meteorological applications (Durran, 2010). Considering a semi-discretization in time the choice of spatial discretization is left open for the moment.

Implicit trapezoidal rule

In this discretization, the governing equations (3.1) are integrated in time from t^n to t^{n+1}, and the time integral on the right-hand side is approximated by the trapezoidal quadrature rule, so that

$$h^{n+1} - h^n = -\frac{H_0 \Delta t}{2} \left(\frac{\partial u^n}{\partial x} + \frac{\partial u^{n+1}}{\partial x} \right) + \Delta t\, b_t^{n+1/2} \,,$$

$$u^{n+1} - u^n = -\frac{\Delta t}{2\mathsf{Fr}^2} \left(\frac{\partial h^n}{\partial x} + \frac{\partial h^{n+1}}{\partial x} \right) \,.$$

Note that it is not specified how to compute the source term at this stage. The notation $b_t^{n+1/2}$ just indicates at which time level the local time derivative of the bottom topography needs to be evaluated to get a second-order accurate approximation. To compute the height h at the new time step one has to solve the (uncritical) Helmholtz problem

$$h^{n+1} - \frac{H_0 \Delta t^2}{4\mathsf{Fr}^2} \frac{\partial^2 h^{n+1}}{\partial x^2} = h^n - H_0 \Delta t \frac{\partial u^n}{\partial x} + \frac{H_0 \Delta t^2}{4\mathsf{Fr}^2} \frac{\partial^2 h^n}{\partial x^2} + \Delta t\, b_t^{n+1/2} \,. \qquad (3.2)$$

The update for u is then given by

$$u^{n+1} = u^n - \frac{\Delta t}{2\mathsf{Fr}^2} \left(\frac{\partial h^n}{\partial x} + \frac{\partial h^{n+1}}{\partial x} \right) \,.$$

The trapezoidal rule is a one-step method, since it only incorporates known values from the previous time step in the computation of the new time level value. Furthermore, it is *A*-stable, which means that the stability region covers the whole complex half plane with negative real parts, as can be seen in Figure 3.1 (left). If the equations to be discretized are linear, the trapezoidal rule is equivalent to the implicit midpoint rule, the latter being a Gauß method which preserves quadratic invariants (Hairer *et al.*, 2006). Thus, for vanishing source term, $b_t \equiv 0$, the trapezoidal rule also preserves the pseudo-energy from (2.11), which is such a quadratic invariant.
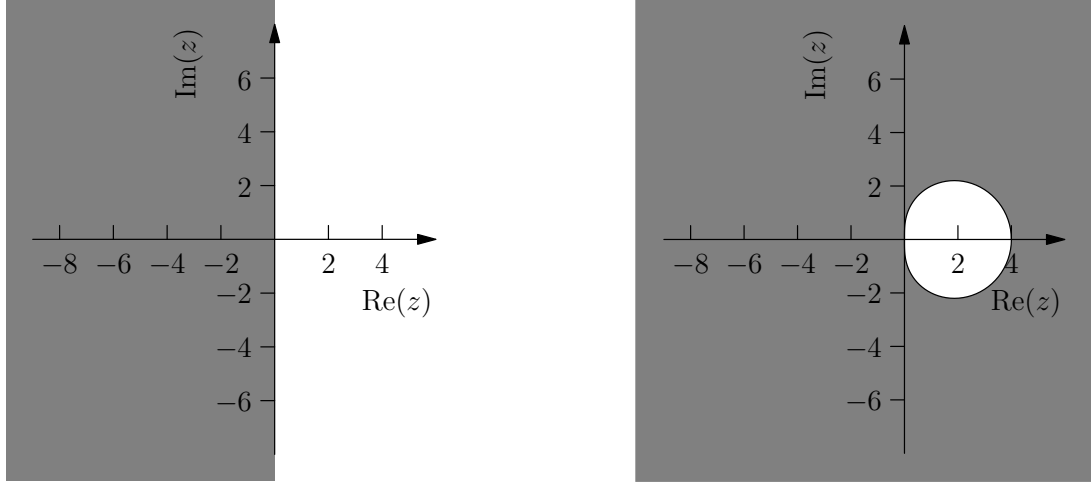
Figure 3.1: Stability regions (shaded) of the implicit trapezoidal rule (left) and BDF(2) scheme (right).

BDF(2) rule

The BDF(2) scheme is one of the simplest methods of the so-called *backward differentiation formulas* (BDF). It is a two-step method, in which the left-hand side is approximated by the derivative of a parabola at t^{n+1}, which interpolates the solution at times t^{n-1}, t^n and t^{n+1}. The discretization of equation (3.1) is then given by

$$\frac{\frac{3}{2}h^{n+1} - 2h^n + \frac{1}{2}h^{n-1}}{\Delta t} = -H_0 \frac{\partial u^{n+1}}{\partial x} + b_t^{n+1} \ ,$$

$$\frac{\frac{3}{2}u^{n+1} - 2u^n + \frac{1}{2}u^{n-1}}{\Delta t} = -\frac{1}{\mathsf{Fr}^2} \frac{\partial h^{n+1}}{\partial x} \ .$$

Rearranging terms, h at the new time step is obtained by the solution of the elliptic problem

$$h^{n+1} - \frac{4H_0\Delta t^2}{9\mathsf{Fr}^2} \frac{\partial^2 h^{n+1}}{\partial x^2} = \frac{4}{3}h^n - \frac{1}{3}h^{n-1} - \frac{2H_0\Delta t}{9}\left(4\frac{\partial u^n}{\partial x} - \frac{\partial u^{n-1}}{\partial x}\right) + \frac{2}{3}\Delta t\, b_t^{n+1} \ , \quad (3.3)$$

and the update for u is given by

$$u^{n+1} = \frac{4}{3}u^n - \frac{1}{3}u^{n-1} - \frac{2\Delta t}{3\mathsf{Fr}^2}\frac{\partial h^{n+1}}{\partial x} \ . \tag{3.4}$$

Being a *L*-stable method, the BDF(2) scheme is a very attractive choice for stiff systems of ordinary differential equations. The stability region of this method is displayed in Figure 3.1 (right).

3.1.2 Super-implicit scheme (extreme BDF)

As can be seen in the result of the asymptotic analysis of the governing equations (2.15), the time derivative of the height becomes singular in the limit $\mathsf{Fr} \to 0$. Thus, in the following scheme the height equation is discretized by a so-called super-implicit scheme. Super-implicit methods are of more implicitness than the so-called implicit formulas in the sense that the approximation of the (highest) time derivative in the equation does not involve values at the new time level, i.e., only the approximation of the right-hand side involves new time level evaluations. They were first introduced by Fukushima (1999) for the application in celestial mechanics.

With the super-implicit discretization of the height equation as it is done in the following, we achieve one crucial property that led us to consider these schemes in the first place: since the time derivative is discretized backwards in time based on the already known height data, the height equation effectively becomes a Poisson equation with two source terms. The first one results from the time derivative of the height and represents (in the context of linear acoustics) the non-acoustic effects of compressibility with second-order accuracy. The second source term is due to the divergence of the velocity flux and appears in the same form also in the pressure projection equation for incompressible flows (see e.g., Schneider *et al.*, 1999; Vater and Klein, 2009; Klein, 2009, and also equation (4.12) in the next chapter). This discretization therefore allows us to make an immediate connection to incompressible or, more generally, sound-proof flow solvers.

For the new discretization the height evolution is approximated by a parabola at t^{n+1} as in the BDF(2) scheme, but this time the interpolation points are t^{n-2}, t^{n-1} and t^n, which leads to

$$\frac{\frac{5}{2}h^n - 4h^{n-1} + \frac{3}{2}h^{n-2}}{\Delta t} = -H_0 \frac{\partial u^{n+1}}{\partial x} + b_t^{n+1} \ .$$
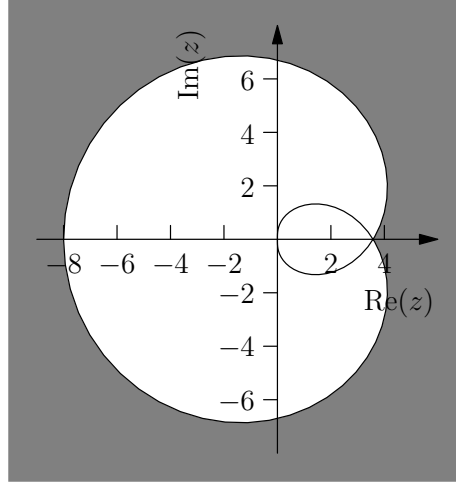
Figure 3.2: Stability region (shaded) and root locus of the super-implicit scheme.

The velocity equation is discretized as in the BDF(2) scheme. The combination of the two equations yields

$$
-\frac{2H_0\Delta t^2}{3\mathsf{Fr}^2}\frac{\partial^2 h^{n+1}}{\partial x^2} = -\left(\frac{5}{2}h^n - 4h^{n-1} + \frac{3}{2}h^{n-2}\right) - \frac{H_0\Delta t}{3}\left(4\frac{\partial u^n}{\partial x} - \frac{\partial u^{n-1}}{\partial x}\right) + \Delta t\, b_t^{n+1}\ .
\tag{3.5}
$$

Thus, this discretization results in the solution of an elliptic equation as ordinary implicit schemes do, but this time it is a Poisson equation instead of a Helmholtz equation for the height at the new time step. The update for the velocity is again (3.4).

Note, however, that this discretization has an "inverse stability constraint" in that it becomes unstable for too small time steps with respect to a fixed mode (see Section 3.1.4 below). This can be also seen from the stability region of this method, which is shown in Figure 3.2 together with the root locus.

3.1.3 Space discretization

As stated above, the space discretization is done using a staggered grid in this study. On this grid the height variables are node centered, i.e. $h_{j+1/2}$, $j = 0, 1, \ldots, M$, where M is the number of cells. The velocity variables, on the other hand, are cell centered, i.e. u_j, $j = 1, 2, \ldots, M$. Standard approximations for the first derivatives of the

height and velocity are given in this context by

$$\frac{\partial h}{\partial x}\bigg|_{x_j} \approx \frac{h_{j+1/2} - h_{j-1/2}}{\Delta x} \quad \text{and} \quad \frac{\partial u}{\partial x}\bigg|_{x_{j+1/2}} \approx \frac{u_{j+1} - u_j}{\Delta x} \ .$$

The corresponding approximation for the second derivative of the height is then given by

$$\frac{\partial^2 h}{\partial x^2}\bigg|_{x_{j+1/2}} \approx \frac{h_{j+3/2} - 2h_{j+1/2} + h_{j-1/2}}{\Delta x^2} \ .$$

Using these discretizations, the elliptic problems as in equations (3.2), (3.3) and (3.5) result in linear systems for the unknown $(h_{j+1/2})$. These systems are then solved using a conjugate gradient algorithm in the applications described below.

3.1.4 Dispersion relations and balanced modes

To be able to quantify the behavior of the numerical discretizations described above, the *discrete-dispersion relations* of these schemes are investigated in the following. Furthermore, the capability of the schemes to reproduce the balanced modes as described by the asymptotic analysis of Section 2.2 are discussed.

The discrete-dispersion relation of a method is obtained by substituting a traveling wave solution of the form

$$\phi_j^n = \hat{\phi} \exp(\mathrm{i}(kj\Delta x - \omega n\Delta t))$$

into the finite-difference formula and solving for ω (Durran, 2010). By separating the frequency into its real and imaginary parts, $\omega = \omega_\mathrm{r} + \mathrm{i}\omega_\mathrm{i}$, one obtains

$$\phi_j^n = \hat{\phi} \exp(\omega_\mathrm{i} n\Delta t) \exp(\mathrm{i}(kj\Delta x - \omega_\mathrm{r} n\Delta t)) = \hat{\phi} A^n \exp(\mathrm{i}(kj\Delta x - \omega_\mathrm{r} n\Delta t)) \ .$$

For a scalar equation, the computation of the imaginary part of ω is equivalent to a von Neumann stability analysis. The *amplification factor* $A := \exp(\omega_\mathrm{i}\Delta t) > 0$ determines how much the mode grows or dissipates per time step. The real part ω_r, on the other hand, describes the *phase speed error*. For a system of equations the same analysis can be carried out, although, in order to obtain stability of the scheme,

the requirement that the amplification factor be less than or equal to unity is merely a necessary condition in general.

The discrete-dispersion relation of the trapezoidal rule applied to the linearized shallow water equations (3.1) on a staggered grid is given by

$$(1 - \xi)^2 + \left(\text{cfl} \cdot \sin \left(\frac{k \Delta x}{2} \right) \right)^2 (1 + \xi)^2 = 0 \ ,$$

where $\xi := \exp(\mathrm{i} \omega \Delta t)$ and $\text{cfl} = \frac{\sqrt{H_0} \Delta t}{\text{Fr} \Delta x} = \frac{c \Delta t}{\Delta x}$ is the Courant (CFL) number (Courant *et al.*, 1928). Solving for ω one obtains for the real part

$$\omega_\mathrm{r} = \pm \frac{2}{\Delta t} \arctan \left(\text{cfl} \cdot \sin \left(\frac{k \Delta x}{2} \right) \right)$$

and for the amplification factor $A \equiv 1$. Therefore, the frequency depends not only on the wave number as in the continuous case, but it is also a function of the Courant number. The derivation of the discrete-dispersion relation for the BDF(2) scheme follows essentially the same route (see Appendix A.1 for the complete derivation for both schemes).

Figures 3.3 and 3.4 show the discrete-dispersion relations for the trapezoidal and the BDF(2) rules applied to the linearized shallow water equations for different Courant numbers. Both schemes slow down modes at almost all wave numbers. This behavior is amplified the higher the wave number and the higher the Courant number are, with the trapezoidal rule featuring the smaller phase speed error of the two schemes throughout. The *group velocity*, which is $\partial \omega / \partial k$ and essentially the speed at which energy is transported, is reduced in a similar manner. For the wave number $\pi / \Delta x$, the highest mode which can be represented on the grid, the group velocity becomes zero. Furthermore, in contrast to the BDF(2) scheme, the trapezoidal scheme is free of numerical dissipation, while in the former one the dissipation is increased the higher the wave number and the higher the Courant number are.

It should be emphasized that the slowdown of high (gravity wave) modes is a common property of implicit methods. By looking at the limit $\text{cfl} \to 0$ both methods,
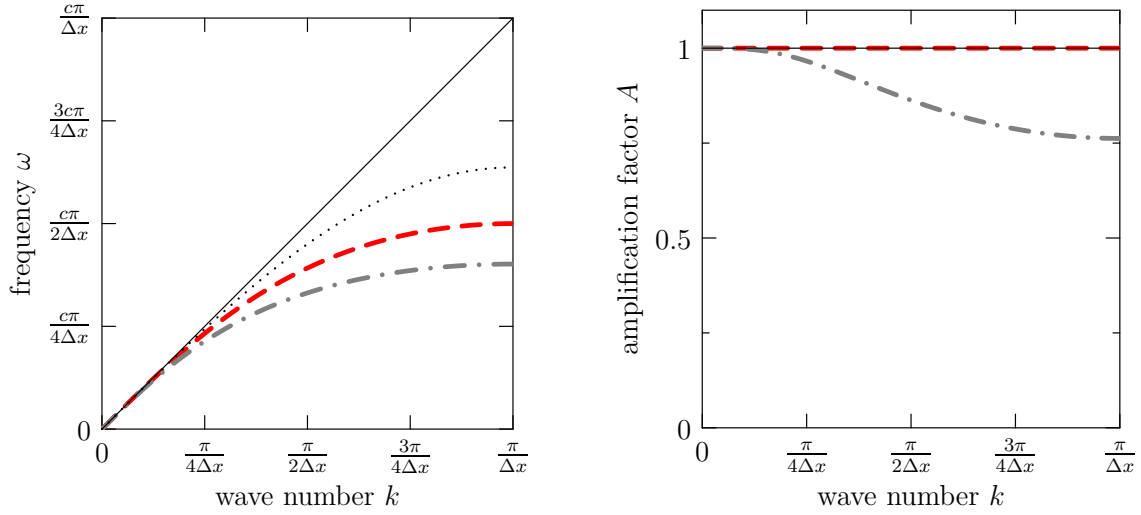
Figure 3.3: Discrete-dispersion relations and amplification factors for the trapezoidal (dashed) and the BDF(2) rules (dot-dashed) applied to the linearized shallow water equations using $\mathrm{cfl} = 1$. The dispersion relation for the continuous system is displayed as black line, and the discrete-dispersion relation for the limiting case $\mathrm{cfl} \to 0$ is given as dotted line.
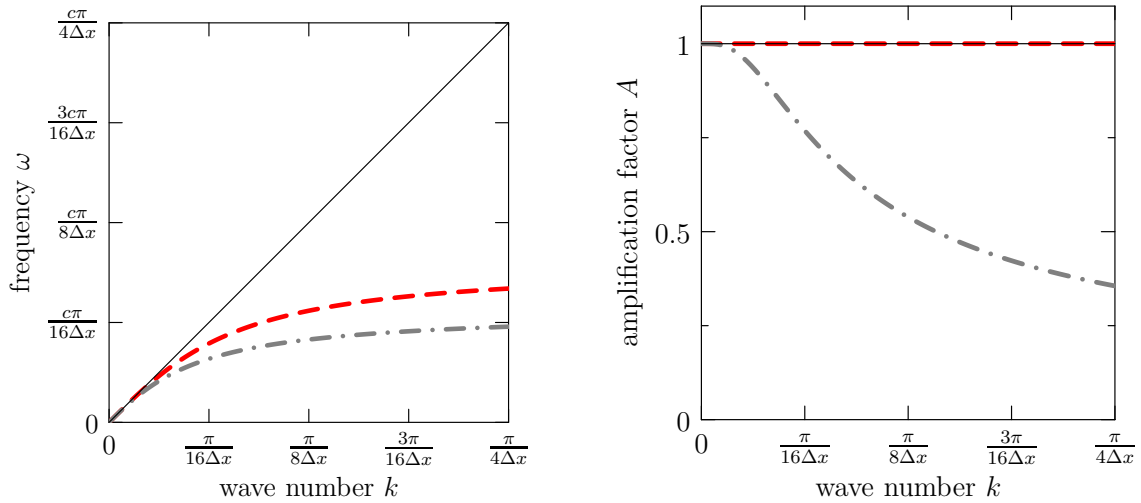

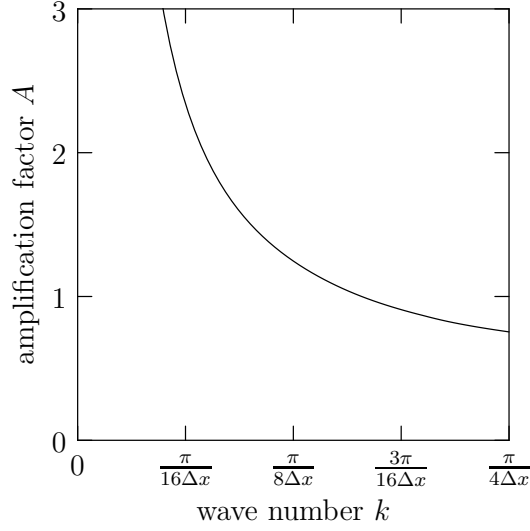**Figure 3.4:** Same as Figure 3.3, but using $\mathrm{cfl} = 10$.

Figure 3.5: Estimated amplification factor for the super-implicit scheme applied to the linearized shallow water equations using $\text{cfl} = 10$.

the trapezoidal rule and the $\text{BDF}(2)$ scheme, result in the dispersion relation

$$\omega_{\mathrm{r}}^{\text{lim}} = \pm \frac{2c}{\Delta x} \sin\left(\frac{k\Delta x}{2}\right) \ .$$

As one can see in Figure 3.3, where the limiting case is plotted with a dotted line, only modes with a wavelength of about 8 grid spacings (corresponds to the wave number $\pi/(4\Delta x)$) and more are advected at a reasonable speed.

For the super-implicit scheme, an analytical expression for the discrete-dispersion relation could not be found. The amplification factor obtained by numerical simulations is displayed in Figure 3.5 for $\text{cfl} = 10$. The graph illustrates the inverse stability behavior of this scheme in that the amplification factor for small wave numbers is greater than 1 at a given time step size. For smaller time steps (resp. Courant numbers) the amplification factor further increases, making the scheme more unstable. The figure further shows, however, that modes with sufficiently high wave numbers are damped.

Concerning the balanced mode, it is of interest how accurately the various schemes maintain the asymptotic balance (2.13) for the case of slow short-wave forcing and whether they are capable of regaining the balance after small initial perturbations of

the system. Assuming that the bottom topography has no contribution on the large scale (i.e. $\bar{b} \equiv 0$) and that there are no gravity waves present arising from initial or boundary conditions, the balance is described by

$$H_0 u_x = b_t \left(t, x \right) \quad \text{and} \quad h \equiv 0$$

up to small perturbations introduced by the variation of the source term in time. The capability to regain the balanced state should be possible for finite time steps, where it is successively approached over a number of time steps. Furthermore, the scheme should reproduce the balanced state in one time step for $\Delta t \to \infty$. These conditions are related to the concepts of A- and L-stability in the mathematical literature of numerical methods for stiff problems (see, e.g., Deuflhard and Bornemann, 2002). However, in the present case the equations are only of oscillatory nature and are not stiff in the classical sense.

The first property, the ability of the method to relax to the balanced state successively, manifests itself through the amplification factor from the stability analysis given above. If the amplification factor is less than unity, the scheme damps out a sudden perturbation and relaxes back to the balanced state. This is true for the BDF(2) scheme, whereas the trapezoidal rule has for all wave numbers an amplification factor of unity. The latter means that any short-wave perturbation, which the scheme interprets as a gravity wave mode, will maintain its amplitude subsequently.

Considering the limit $\Delta t \to \infty$, one obtains for the trapezoidal rule the relations

$$\frac{\partial h^{n+1}}{\partial x} = -\frac{\partial h^n}{\partial x} \quad \text{and} \quad H_0 \frac{\partial u^{n+1}}{\partial x} = -H_0 \frac{\partial u^n}{\partial x} + 2b_t^{n+1/2} \ .$$

This reflects the behavior described above that any perturbation cannot dissipate. The first derivatives of height and velocity essentially oscillate around the balanced state. In case of the BDF(2) and the super-implicit schemes the situation is quite different. In the limit $\Delta t \to \infty$ one obtains for both methods

$$\frac{\partial h^{n+1}}{\partial x} = 0 \quad \text{and} \quad H_0 \frac{\partial u^{n+1}}{\partial x} = b_t^{n+1} \ .$$

Thus, the schemes achieve balance in a single, sufficiently large, time step. This behav-

ior is characteristic of backward differentiation formulas by construction (Deuflhard and Bornemann, 2002).

As a consequence, the practitioner is faced with the following problem: On the one hand, one would like to minimize dispersion and preserve the amplitude of well resolved modes. For this purpose the trapezoidal rule seems to be the best suited one of the described schemes. On the other hand, the solution should rapidly relax to the balanced mode in case of short wave number forcing. This property is better reproduced by the backward differentiation formulas. In the next section, a strategy is presented for combining the two aspects into one single, scale-dependent numerical time integrator.

3.2 Multilevel method for long-wave linearized shallow water flows

As indicated earlier, by "long-wave" we denote here highly resolved gravity wave modes that oscillate very slowly in comparison with the shortest modes that could potentially be represented on the grid. The goal is to provide a discretization that guarantees the following: given a time step, the scheme will automatically filter all gravity wave modes from the initial data that have characteristic frequencies which are not resolved in time, whereas all sufficiently long waves with lower frequencies get to be resolved and accurately computed. Thus the intent is to avoid the dichotomy between either damping all gravity wave modes or accepting spurious, slowly moving short-wave modes as outlined in the introduction.

3.2.1 General idea

Let us assume that there exist direct scale dependent splittings of the height and velocity fields, i.e.,

$$h = \sum_{\nu=0}^{\nu_M} h^{(\nu)} \quad \text{and} \quad u = \sum_{\nu=0}^{\nu_M} u^{(\nu)} \; . \tag{3.6}$$

Ideally, this could be a quasi-spectral or wavelet decomposition, splitting h and u into (local) high wave number and low wave number components. However, these splittings are still to be determined, and it will be seen that they cannot be defined independently in order to get optimal results. By the introduction of projection operators Π_ν^h and Π_ν^u, which project a height or velocity field to the scale ν, these splitting parts shall be given by

$$h^{(\nu)} = (\Pi_\nu^h - \Pi_{\nu-1}^h)h \quad \text{and} \quad u^{(\nu)} = (\Pi_\nu^u - \Pi_{\nu-1}^u)u \; , \tag{3.7}$$

where the auxiliary operators $\Pi_{-1}^h \equiv 0$ and $\Pi_{-1}^u \equiv 0$ have been defined, for simplicity.

Furthermore, let us consider two different time discretizations of the equations for linearized shallow water flows, which are linear in h^{n+1} and u^{n+1}. In general, these can be written as

$$\begin{aligned}
A_1 h^{n+1} &= f_1^h(h^n, u^n, h^{n-1}, u^{n-1}, \dots) \; , \\
u^{n+1} &= f_1^u(h^{n+1}, h^n, u^n, h^{n-1}, u^{n-1}, \dots) \; ,
\end{aligned} \tag{3.8}$$

and

$$\begin{aligned}
A_2 h^{n+1} &= f_2^h(h^n, u^n, h^{n-1}, u^{n-1}, \dots) \; , \\
u^{n+1} &= f_2^u(h^{n+1}, h^n, u^n, h^{n-1}, u^{n-1}, \dots) \; ,
\end{aligned} \tag{3.9}$$

where the equations incorporating the linear operators A_i represent the discretizations of the elliptic equations to obtain the new height. The second equation of (3.8) and (3.9), respectively, is the update for the velocity equation. A convex scale-wise combination of the two schemes with scale dependent weights μ_ν results in the scheme

$$\begin{aligned}
\sum_{\nu=0}^{\nu_M} \left(\mu_\nu (A_1 h^{n+1})^{(\nu)} + (1 - \mu_\nu)(A_2 h^{n+1})^{(\nu)} \right) &= \sum_{\nu=0}^{\nu_M} \left(\mu_\nu f_1^{h,(\nu)} + (1 - \mu_\nu) f_2^{h,(\nu)} \right) \; , \\
u^{n+1} &= \sum_{\nu=0}^{\nu_M} \left(\mu_\nu f_1^{u,(\nu)} + (1 - \mu_\nu) f_2^{u,(\nu)} \right) \; .
\end{aligned} \tag{3.10}$$

Here, each term with an index (ν) is the part associated to the scale ν through (3.7).

Since A_1 and A_2 are linear, a new scale dependent operator A can be defined by

$$A := \sum_{\nu=0}^{\nu_M} \left(\mu_\nu A_1 + (1 - \mu_\nu)A_2\right) \left(\Pi_\nu^h - \Pi_{\nu-1}^h\right) .$$

With this definition the first equation of (3.10) becomes

$$A h^{n+1} = \sum_{\nu=0}^{\nu_M} \left(\mu_\nu f_1^{h,(\nu)} + (1 - \mu_\nu)f_2^{h,(\nu)}\right) .$$

For implicit discretizations, such as the standard time integration schemes described above this again results in an (uncritical) Helmholtz problem, but this time with a scale dependent operator A.

As a first example consider the blending of the implicit trapezoidal rule with the BDF(2) scheme. For these schemes, the linear operators for the determination of the new time level height are given by

$$A_{\text{TRA}} = \text{id} - \frac{H_0 \Delta t^2}{4 \text{Fr}^2} \partial_{xx} \quad \text{and} \quad A_{\text{BDF2}} = \text{id} - \frac{4 H_0 \Delta t^2}{9 \text{Fr}^2} \partial_{xx} .$$

Taking the μ-dependent convex combination of the two operators and summing over the scales results in

$$h^{n+1} - \frac{H_0 \Delta t^2}{\text{Fr}^2} \sum_{\nu=0}^{\nu_M} \left(\frac{\mu_\nu}{4} + \frac{4(1 - \mu_\nu)}{9}\right) h_{xx}^{(\nu),n+1} = \\ \sum_{\nu=0}^{\nu_M} \left(\mu_\nu f_{\text{TRA}}^{h,(\nu)} + (1 - \mu_\nu)f_{\text{BDF2}}^{h,(\nu)}\right) , \tag{3.11}$$

where

$$f_{\text{TRA}}^{h,(\nu)} = \left[h^n - H_0 \Delta t \, u_x^n + \frac{H_0 \Delta t^2}{4 \text{Fr}^2} h_{xx}^n + \Delta t \, b_t^{n+1/2}\right]^{(\nu)} ,$$

$$f_{\text{BDF2}}^{h,(\nu)} = \left[\frac{4}{3}h^n - \frac{1}{3}h^{n-1} - \frac{2 H_0 \Delta t}{9}\left(4u_x^n - u_x^{n-1}\right) + \frac{2}{3}\Delta t \, b_t^{n+1}\right]^{(\nu)} .$$

The velocity update reads

$$u^{n+1} = \sum_{\nu=0}^{\nu_M} \mu_\nu \left[u^n - \frac{\Delta t}{2\mathsf{Fr}^2} \left(h_x^n + h_x^{n+1} \right) \right]^{(\nu)} +$$
$$(1 - \mu_\nu) \left[\frac{4}{3} u^n - \frac{1}{3} u^{n-1} - \frac{2\Delta t}{3\mathsf{Fr}^2} h_x^{n+1} \right]^{(\nu)} . \tag{3.12}$$

Note that the discretization of the source term b_t is not specified at this point, and it is assumed that this term could be exactly integrated.

In a second example, the trapezoidal rule is combined with the super-implicit (extreme BDF) scheme described in Section 3.1.2. The resulting update is given by

$$\sum_{\nu=0}^{\nu_M} \mu_\nu h^{(\nu),n+1} - \frac{H_0 \Delta t^2}{4\mathsf{Fr}^2} h_{xx}^{n+1} = \sum_{\nu=0}^{\nu_M} \left(\mu_\nu f_{\mathrm{TRA}}^{h,(\nu)} + (1 - \mu_\nu) f_{\mathrm{EBDF}}^{h,(\nu)} \right) , \tag{3.13}$$

where

$$f_{\mathrm{TRA}}^{h,(\nu)} = \left[h^{(\nu),n} - H_0 \Delta t\, u_x^{(\nu),n} + \frac{H_0 \Delta t^2}{4\mathsf{Fr}^2} h_{xx}^{(\nu),n} + \Delta t\, b_t^{(\nu),n+1/2} \right]^{(\nu)} ,$$

$$f_{\mathrm{EBDF}}^{h,(\nu)} = \frac{3}{8} \left[-\frac{5}{2} h^{(\nu),n} + 4 h^{(\nu),n-1} - \frac{3}{2} h^{(\nu),n-2} \right.$$
$$\left. - \frac{H_0 \Delta t}{3} \left(4 u_x^{(\nu),n} - u_x^{(\nu),n-1} \right) + \Delta t\, b_t^{(\nu),n+1} \right]^{(\nu)} .$$

The update for the velocity is again given by (3.12).

The main difference between the two variants of the method is that in the first case the difference operator acting on the height variables incorporates the multiscale information in its discrete Laplacian part (second term of LHS in equation (3.11)), whereas in the second case it appears in the Helmholtz part (first term of LHS in equation (3.13)). Note, that if the scale dependent splitting (3.6) does not diagonalize the differential operator, the Helmholtz problem (3.11) (resp. (3.13)) must be solved as a whole, and the solution components on the different scales are coupled to each other.
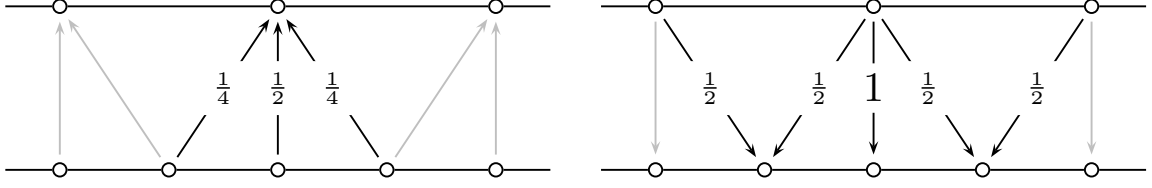
Figure 3.6: One-dimensional versions of full weighting (left) and linear interpolation (right) operators known from standard finite difference geometric multigrid. Arrows indicate mappings between grid functions associated with grid nodes.

3.2.2 Scale splitting

To define the operators in the scale-dependent discretizations, the quasi-spectral decompositions of the height and velocity fields needs to be carefully specified. Furthermore, it will be seen that the two splittings cannot be defined independently, but have to satisfy certain relationships.

First, the decomposition of the height field is discussed, and in a second step the appropriate velocity decomposition is derived. In order to define the grid hierarchy, let us assume that the total number of grid cells in the domain is a power of two. A coarser grid is then obtained by eliminating every second grid node or, equivalently, by merging two adjacent cells. In the current approach the restriction and prolongation operators used in standard multigrid algorithms are utilized to define the space decomposition. One of the most commonly used operator pairs for finite difference approximations (and in a slightly modified way also for finite element approximations) is the *full weighting* (restriction) and the *linear interpolation* (prolongation) (Trottenberg *et al.*, 2001). They can be defined by their stencil. The full weighting is given by

$$R^{(\nu)} = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} ,$$

which means that a variable on the coarse grid node at grid level (ν) is derived by averaging over the values at the same node and the two adjacent nodes on the fine grid at grid level $(\nu + 1)$ with the weights given in the stencil above (see also Figure 3.6, left). The linear interpolation from grid level (ν) to grid level $(\nu + 1)$ is given by

$$P^{(\nu)} = \frac{1}{2} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} . \tag{3.14}$$
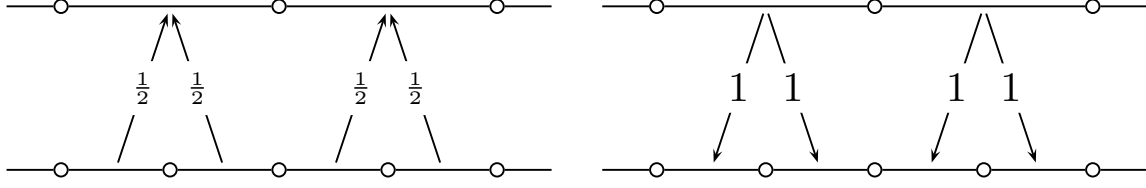
Figure 3.7: One-dimensional versions of restriction (left) and prolongation (right) operators known from standard finite volume geometric multigrid. Arrows indicate mappings between grid functions associated with grid cells (instead of with grid nodes as in Figure 3.6).

This means that the height at grid nodes living on the fine grid level, which have a common coarse grid node, obtain the same value as on the coarse grid. The values at grid nodes in between are computed by the average of the values of the adjacent grid nodes (Figure 3.6, right). Note, that $P^{(\nu)}$ and $R^{(\nu)}$ are adjoint up to a scaling factor. This is the classical way in which these operators are defined.

Other alternatives are also possible. A common restriction/prolongation pair used in finite volume approximations for cell centered variables is given by

$$\tilde{R}^{(\nu)} = \frac{1}{2} \begin{bmatrix} 1 & 1 \end{bmatrix} \quad \text{and} \quad \tilde{P}^{(\nu)} = \begin{bmatrix} 1 & 1 \end{bmatrix} . \tag{3.15}$$

Here, the value for the coarse grid cell is computed by simply averaging over the two corresponding small cells on the fine grid, and the values for the fine grid cells are obtained by just copying the value from the corresponding coarse grid cell (cf. Figure 3.7).

The scale splitting is now defined as follows. Let φ be a grid function, which is decomposed into parts $\varphi^{(\nu)}$ associated to different grid levels with

$$\varphi = \sum_{\nu=0}^{\nu_M} \varphi^{(\nu)} .$$

Then, the grid function on the coarsest level is obtained by the operation

$$\varphi^{(0)} = \left(R^{(0)} \circ R^{(1)} \circ \cdots \circ R^{(\nu_M-1)} \right) \varphi ,$$

and the grid functions on finer levels are computed by

$$\varphi^{(\nu)} = \left(I - P^{(\nu-1)} \circ R^{(\nu-1)} \right) \circ \left(R^{(\nu)} \circ R^{(\nu+1)} \circ \cdots \circ R^{(\nu_M-1)} \right) \varphi \ .$$

On a staggered grid the splitting in the velocity field cannot be the same as the one for the height. Since the velocity variable is cell centered, only transfer operators as in (3.15) are applicable. Ideally, the splitting should be chosen such that only the portion of the height associated with the grid level (ν) enters the update for the velocity on the same grid level. Revisiting equation (3.12) shows that only first derivatives of the height at different time levels enter the update. Therefore, the splitting in the velocity must match the splitting in $\partial h/\partial x$ induced by the h-splitting.

Considering just two grid levels, where the fine grid has nodes $x_{j+1/2}$, $j = \ldots, -1, 0, 1, \ldots$, and the coarse grid has nodes $x_{2j+1/2}$, $j = \ldots, -1, 0, 1, \ldots$ the splitting of the velocity can be derived as follows: Let us denote variables living on the fine grid level by $h^{(f)}_{j+1/2}$ and $u^{(f)}_j$ and those associated with the coarse grid by $h^{(c)}_{2j+1/2}$ and $u^{(c)}_{2j+3/2}$. Then, by enforcing the above relationship between height and velocity on a discrete level leads to

$$
\begin{aligned}
u^{(c)}_{2j+3/2} &= \frac{h^{(c)}_{2j+5/2} - h^{(c)}_{2j+1/2}}{2\Delta x} \\
&= \frac{1}{2\Delta x} \left(\frac{1}{4} h^{(f)}_{2j+7/2} + \frac{1}{2} h^{(f)}_{2j+5/2} + \frac{1}{4} h^{(f)}_{2j+3/2} - \frac{1}{4} h^{(f)}_{2j+3/2} - \frac{1}{2} h^{(f)}_{2j+1/2} - \frac{1}{4} h^{(f)}_{2j-1/2} \right) \\
&= \frac{1}{8\Delta x} \left(h^{(f)}_{2j+7/2} - h^{(f)}_{2j+5/2} \right) + \frac{3}{8\Delta x} \left(h^{(f)}_{2j+5/2} - h^{(f)}_{2j+3/2} \right) + \\
&\quad \frac{3}{8\Delta x} \left(h^{(f)}_{2j+3/2} - h^{(f)}_{2j+1/2} \right) + \frac{1}{8\Delta x} \left(h^{(f)}_{2j+1/2} - h^{(f)}_{2j-1/2} \right) \\
&= \frac{1}{8} u^{(f)}_{2j+3} + \frac{3}{8} u^{(f)}_{2j+2} + \frac{3}{8} u^{(f)}_{2j+1} + \frac{1}{8} u^{(f)}_{2j} \ .
\end{aligned}
$$

Here, the prolongation defined by (3.14) has been used in the second equality to translate coarse grid variables to fine grid variables. This results in a restriction with stencil

$$\hat{R}^{(\nu)} = \frac{1}{8} \begin{bmatrix} 1 & 3 & 3 & 1 \end{bmatrix}$$

for the velocity. The obvious choice for the prolongation operator is a scaled version

of the adjoint of the restriction operator $\hat{R}^{(\nu)}$, which results in

$$\hat{P}^{(\nu)} = \frac{1}{4} \begin{bmatrix} 1 & 3 & 3 & 1 \end{bmatrix} . \tag{3.16}$$

However, although it might introduce some inconsistency, it turns out that in some cases

$$\tilde{P}^{(\nu)} = \begin{bmatrix} 1 & 1 \end{bmatrix} \tag{3.17}$$

is a more stable alternative.

Another issue arises from the aim to have on each grid level essentially the same discretization of the Laplacian. Considering again only two grid levels, where the splitting at each node is given by $h_{j+1/2} = h_{j+1/2}^{(f)} + h_{j+1/2}^{(c)}$, the discrete Laplacian evaluated at nodes common to both grids splits like

$$
\begin{aligned}
L(h)_{2j+1/2} &= \frac{h_{2j-1/2} - 2h_{2j+1/2} + h_{2j+3/2}}{\Delta x^2} \\
&= \frac{h_{2j-1/2}^{(f)} - 2h_{2j+1/2}^{(f)} + h_{2j+3/2}^{(f)}}{\Delta x^2} + \frac{h_{2j-1/2}^{(c)} - 2h_{2j+1/2}^{(c)} + h_{2j+3/2}^{(c)}}{\Delta x^2} \\
&= L(h)_{2j+1/2}^{(f)} + \frac{1}{\Delta x^2} \left(\frac{1}{2} h_{2j-3/2}^{(c)} - h_{2j+1/2}^{(c)} + \frac{1}{2} h_{2j+5/2}^{(c)} \right) \\
&= L(h)_{2j+1/2}^{(f)} + \frac{2}{(2\Delta x)^2} \left(h_{2j-3/2}^{(c)} - 2h_{2j+1/2}^{(c)} + h_{2j+5/2}^{(c)} \right) \\
&= L(h)_{2j+1/2}^{(f)} + 2L(h)_{2j+1/2}^{(c)} .
\end{aligned}
$$

Note, that in the third equality the substitution $h_{2j+3/2}^{(c)} = \frac{1}{2}(h_{2j+1/2}^{(c)} + h_{2j+5/2}^{(c)})$ was used, which denotes linear interpolation between grid points. At the nodes only belonging to the fine grid, the discrete Laplacian is

$$
\begin{aligned}
L(h)_{2j+3/2} &= \frac{h_{2j+1/2}^{(f)} - 2h_{2j+3/2}^{(f)} + h_{2j+5/2}^{(f)}}{\Delta x^2} + \frac{h_{2j+1/2}^{(c)} - 2h_{2j+3/2}^{(c)} + h_{2j+5/2}^{(c)}}{\Delta x^2} \\
&= L(h)_{2j+1/2}^{(f)} + \frac{1}{\Delta x^2} \cdot 0 = L(h)_{2j+1/2}^{(f)} .
\end{aligned}
$$

Therefore, the same discretization for the Laplacian can be applied on each grid level, and in the summation two times the result of the coarse grid level needs to be added to the common nodes on the next finer grid level. This is not common to all splittings:

If the finite volume type splitting defined by (3.15) is chosen, the splitting of the Laplacian into fine and coarse grid portions is given by

$$L(h)_{2j} = L(h)_{2j}^{(f)} + \frac{1}{\Delta x^2} \left(h_{2j-3/2}^{(c)} - h_{2j+1/2}^{(c)} \right)$$

for even fine grid cells and

$$L(h)_{2j+1} = L(h)_{2j+1}^{(f)} + \frac{1}{\Delta x^2} \left(-h_{2j+1/2}^{(c)} + h_{2j+5/2}^{(c)} \right)$$

for odd fine grid cells, where it was used that $h_{2j+1/2}^{(c)} = h_{2j}^{(c)} = h_{2j+1}^{(c)}$ (which means that piecewise constant data are assumed). This means that different discretizations of the Laplacian must be applied on each grid level.

To complete the description of the new scheme, the weighting function μ_ν has still to be defined. Generally, this choice is not restricted besides the requirement that the resulting scheme should be stable. In the current implementation this quantity is given by

$$\mu_\nu = \begin{cases} \min\left(1, \dfrac{\nu_M - \nu}{\lfloor \log_2 \mathrm{cfl} \rfloor} \right) & \text{if cfl} \geq 2 \,, \\ 1 & \text{otherwise} \,, \end{cases} \tag{3.18}$$

where $\lfloor \cdot \rfloor$ means rounding towards minus infinity. Thus, μ_ν is chosen such that the scheme in equation (3.10) associates the standard implicit trapezoidal scheme with all gravity wave modes corresponding to coarse grids with grid-CFL number $\mathrm{cfl}_\nu \leq 1$ ($\mu_\nu = 1$), while the discretization is nudged towards BDF(2) or super-implicit for modes living on grids with $\mathrm{cfl}_\nu > 1$ ($\mu_\nu < 1$). However, if the fine-grid-CFL number is smaller than 2, the scheme would consequently end up with using only the trapezoidal rule.

3.2.3 Dispersion relation

For a concise analysis it would be desirable to derive a discrete dispersion relation and the amplification factor depending on the wave number also for the multilevel schemes. However, this is not possible. The reason lies in the quasi-spectral splitting of the dependent variables. Consider for example a sinusoidal simple wave, which

represents a given wave number. When this initial data is applied to the multilevel schemes, the splitting of the grid functions to the different grid levels does not assign the whole data to one grid level, in general. On the contrary, since the splitting should not be spectral but rather have a local character, a Fourier mode is distributed to different levels. Assuming varying blending weights for the levels, each fraction of the data associated to one grid level is processed with a different scheme having its own dispersion relation. Therefore, a discrete-dispersion relation as it is derived for the single-level schemes cannot be obtained.

Ideally, all traveling wave solutions would have to be found, which translate by a certain amount in one step of the multilevel scheme and only vary in amplitude. This is related to an eigenvalue analysis, where the matrix is analyzed, which translates the solution vector $(h^n, u^n, h^{n-1}, u^{n-1})$ into $(h^{n+1}, u^{n+1}, h^n, u^n)$ etc. In this work, this path is not further pursued. Instead, a discrete-dispersion relation and amplification factor is estimated by numerical simulation with the common Fourier modes as initial data. The result for the multilevel trapezoidal/BDF(2) scheme at CFL number cfl = 10 is given in Figure 3.8, where for comparison the discrete-dispersion relations for the trapezoidal rule and the BDF(2) scheme are also plotted. The multilevel scheme is applied with five grid levels and blending weights $\mu_{\boldsymbol{\nu}} = (1, 1, 2/3, 1/3, 0)$. The graph shows for low wave numbers that the dispersion and diffusion errors are close to those of the trapezoidal rule. Moreover, for higher wave numbers the dispersion behavior resembles more the one of the BDF(2) scheme. At wave numbers around $3/(16\Delta x)$, which is equivalent to a wave length of about 10 grid spacings, and higher the fact that the Fourier modes are no traveling wave solutions of the scheme becomes apparent. But these modes are all damped as suggested by the computed amplification factor. However, the results for dispersion relation and amplification factor should be treated with some caution for these modes.

3.2.4 Interaction between time and space discretizations

To further analyze the properties of the Helmholtz operator associated with the new scheme the corresponding matrix was assembled. By applying the operator to the unit vector $e_i = (\delta_{ij})_{j=1,\dots,n}$ the i-th column of the matrix is obtained. In Figure 3.9, the sparsity pattern for the resulting matrices of the multilevel trapezoidal/BDF(2) scheme
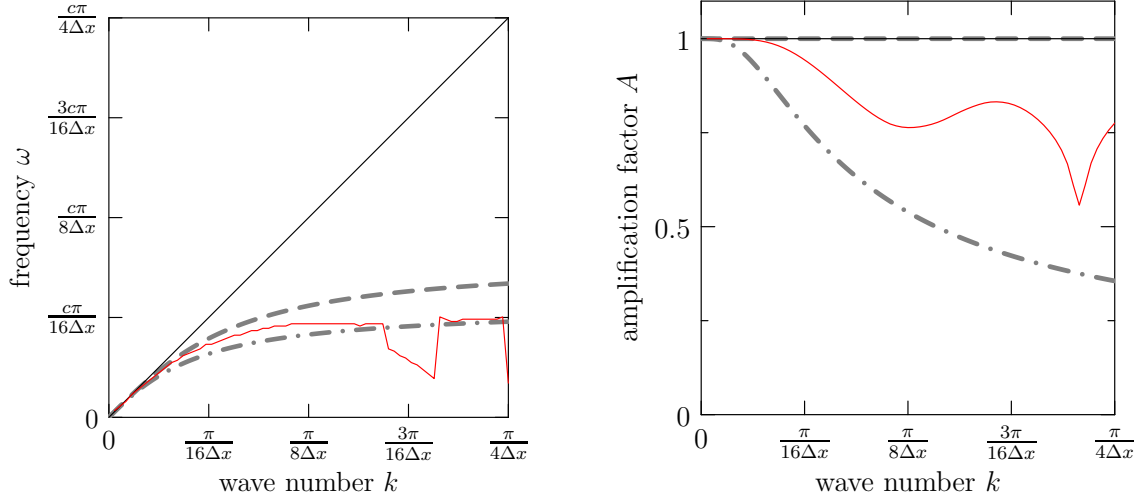
Figure 3.8: Estimated discrete-dispersion relation and amplification factor for the multilevel trapezoidal/BDF(2) scheme (red) applied to the linearized shallow water equations using $\mathrm{cfl} = 10$. For comparison the dispersion relations for the continuous system (black), the trapezoidal rule (gray dashed) and the BDF(2) scheme (gray dot-dashed) are displayed.

(left) and the multilevel trapezoidal/super-implicit scheme (right) are displayed for a domain with 256 grid cells. The results are obtained by using five grid levels and μ_ν changing linearly from the coarsest level ($\mu_0 = 1$) to the finest level ($\mu_4 = 0$). Obviously, the matrices do not have the same typical tridiagonal pattern as the blended base schemes. In contrast, the different time integration schemes for each grid level influence the effective space discretization. The differences in the pattern between the two versions of the new scheme is essentially due to the fact that in case of the multilevel trapezoidal/BDF(2) scheme the multiscale information is hidden in the Laplacian part of the operator, and in case of the multilevel trapezoidal/super-implicit scheme it is hidden in the Helmholtz part.

It should be emphasized that the band width of the matrix strongly depends on the number of grid levels and the blending parameter μ_ν. Further off-diagonal entries only appear if the blending weight changes between successively coarser grid levels. Thus, the given choice displays the maximum with respect to non-zero matrix entries for this number of levels. For μ_ν as defined in (3.18) and a Courant number $\mathrm{cfl} = 10$,
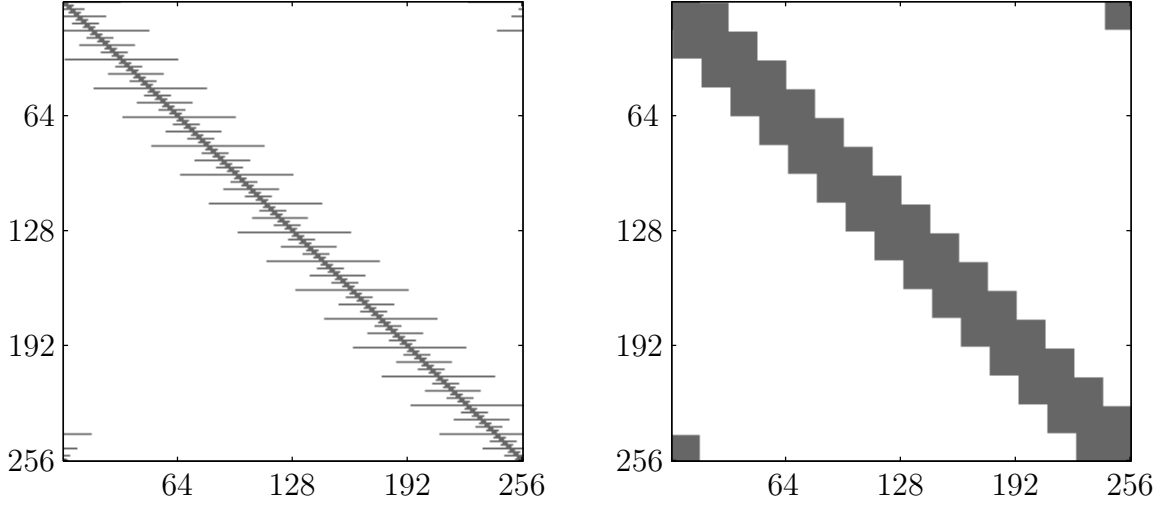
Figure 3.9: Sparsity pattern for the resulting matrices of multilevel schemes, domain with 256 grid cells and five grid levels. Left: trapezoidal/BDF(2); right: trapezoidal/super-implicit.

the band width of the matrix is slightly reduced. The deviations in the matrix entries from those resulting from the single-scale fine grid time integrator are relatively small. Compared to the maximum absolute value appearing in the matrix they have a relative order of magnitude of at most four percent in case of the diagonal elements and of at most one percent for the off-diagonal elements. However, these estimates also depend on the number of grid levels and the blending parameter.

The resulting elliptic problem with such an operator can easily be solved with any iterative method for the solution of linear systems, where the operator is only applied to grid functions in each iteration step. Popular representatives of this class are Krylov subspace methods. But this approach might be too inefficient when considering problems in more than one space dimension. Therefore, new algorithms might have to be proposed that, e.g., borrow ideas from multigrid methods for elliptic problems.

3.3 Numerical results

The new scheme is applied to two test cases, and the results are compared to those of the standard reference methods. The first test case explores the desired property to treat freely propagating short-wave data, which is not well resolved in time, differently compared to well resolved long-wave data. For this purpose, the scheme is initialized with "multiscale" initial data and no source term is present. In a second test case, the ability of the scheme to relax to a balanced state is analyzed. In order to do this, a source term with the appropriate scaling is introduced. For the new multilevel scheme, always five grid levels are used in the considered test cases.

3.3.1 Multiscale initial data

In this test case the previously described numerical schemes are compared by applying them to either pure long-wave initial data, or "multiscale" initial data in a periodic domain $\xi \in [0, 1]$. Here, the governing equations are transformed to the large scale variable ξ, in which the gravity wave speed $c_\xi = \sqrt{H_0}$ is of order 1. In particular, $H_0 = 1$ for this test case. The relationship between height and velocity is chosen in such a way that a right running (gravity) simple wave is obtained. The pure long-wave initial data is given by

$$u(0, \xi) = u_0(\xi - \xi_0) \quad \text{and} \quad h(0, \xi) = \mathsf{Fr}\,\sqrt{H_0}\,u(0, \xi) \; ,$$

where

$$u_0(\xi) = \exp\left(-\left(\frac{\xi}{\sigma_0}\right)^2\right)$$

with $\xi_0 = 0.75$ and $\sigma_0 = 0.1$, and the "multiscale" initial data is

$$u(0, \xi) = u_0(\xi - \xi_0) + u_1(\xi - \xi_1) \quad \text{and} \quad h(0, \xi) = \mathsf{Fr}\,\sqrt{H_0}\,u(0, \xi) \; ,$$

where u_0 is defined as above and

$$u_1(\xi) = u_0(\xi)\cos\left(\frac{k\xi}{\mathsf{Fr}\,\sigma_0}\right) = u_0(\xi)\cos\left(\frac{kx}{\sigma_0}\right)$$
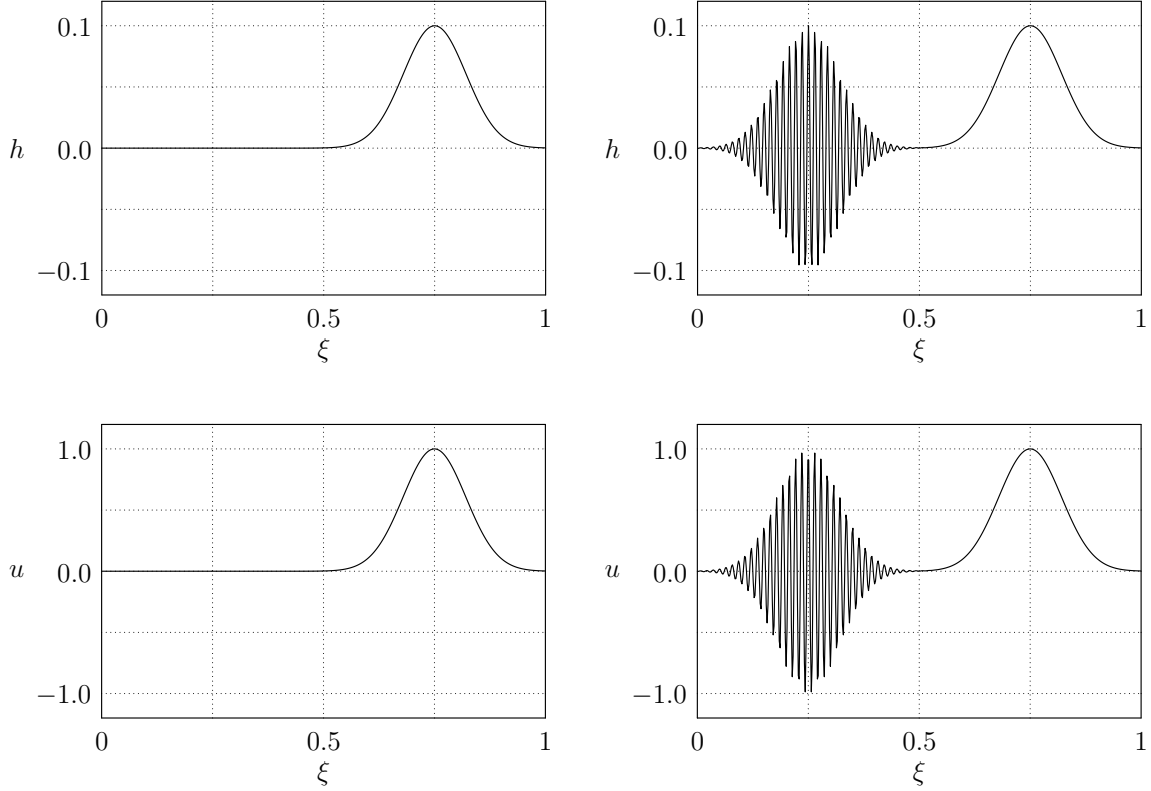
Figure 3.10: Initial conditions for the simple wave test case with single-scale pure long-wave data (left) and "multiscale" data (right). Top row: height, bottom row: velocity.

with $\xi_1 = 0.25$ and $k = 0.7 \cdot 2\pi$ (see Figure 3.10). The Froude number is set to $\mathsf{Fr} = 0.1$. Note, that the correct scaling of h and u is already given by the initial data. No source term is present, i.e., $b \equiv 0$.

The schemes are applied to this initial data on a grid with 512 cells (i.e., $\Delta\xi = 1/512$) and a Courant number $\mathrm{cfl}_{\Delta\xi} = 10$. They are compared at a final time $t_{\mathrm{end}} = 3$, which is equivalent to 154 time steps. At this time the exact solution is identical to the initial data, and the wave has traveled three times across the domain. The blending weights for the multilevel schemes are given by $\mu_{\boldsymbol{\nu}} = (1, 1, 2/3, 1/3, 0)$.

The implicit trapezoidal rule described in Section 3.1.1 produces the results in Figure 3.11. Here, and in the following, only height is displayed, whenever the velocity field is essentially the same. The results show what has already been revealed theoretically by the discrete-dispersion relation for large time steps: the scheme
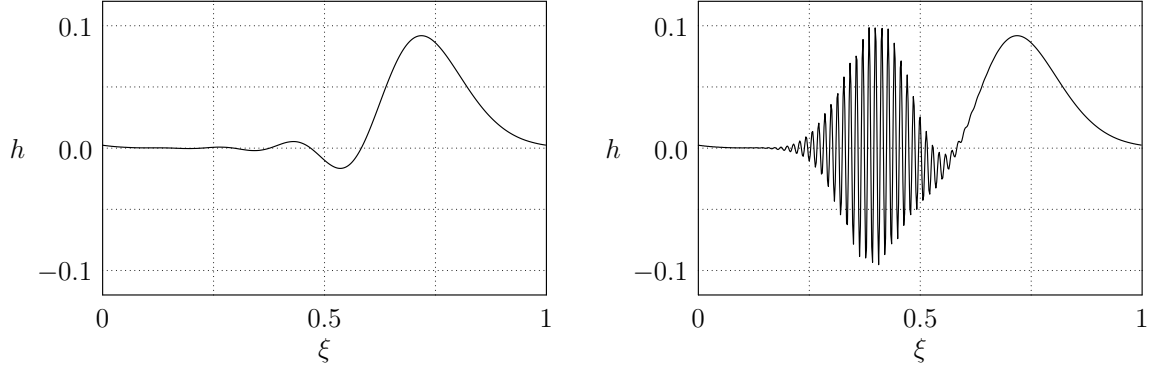
Figure 3.11: Numerical solution (height) of the simple wave test case using the trape-zoidal rule on a grid with 512 cells and $\mathrm{cfl} = 10$ at time $t_{\mathrm{end}} = 3$. Left: results obtained with single scale initial data; right: results obtained with "multiscale" initial data.

achieves large-CFL stability by slowing down the short-wave components of the solution. In the single-scale example, the consequence of this numerical dispersion error is a slight distortion of the Gaussian pulse which generates new artificial extrema, and slightly slows down the wave. When run over longer times, this trend continues and the Gaussian pulse decomposes into an essentially uncorrelated superposition of Fourier modes of various length scales.

For the multiscale initial data, the resultant error is much more dramatic: the numerical scheme does not only slow down the phase speed of the short-wave components of the solution, but also the group velocity drops almost to zero for these components. As a consequence, the long-wave pulse in Figure 3.11 (right) has passed the domain three times as it should, but the short-wave oscillations have essentially stayed in place. Furthermore, their amplitude has not diminished.

Clearly, such a behavior, when extrapolated to global atmospheric flow simulations, would be unacceptable. Not only would the planetary scale Lamb waves be slightly distorted, but any small-scale divergence induced by one of the ubiquitous diabatic source terms in an atmosphere model would potentially set up stationary, short-wave, fake numerical standing waves (see also the results of the second test case). In interaction with the parametrizations of various physical subgrid-scale processes, nota bene the physics of moisture, these short-wave modes will likely produce non-trivial erroneous consequences for the further flow evolution.
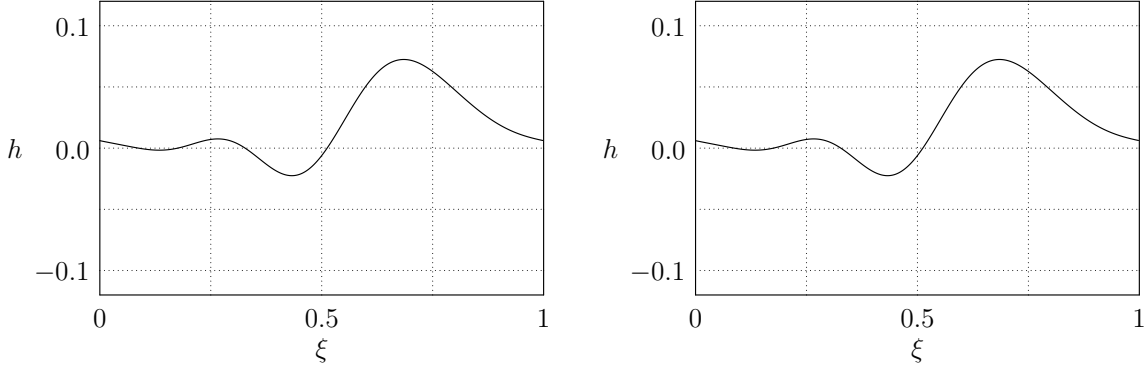
Figure 3.12: Numerical solution (height) of the simple wave test case using the BDF(2) scheme on a grid with 512 cells and cfl = 10 at time $t_{\mathrm{end}} = 3$. Left: results obtained with single scale initial data; right: results obtained with "multiscale" initial data.

The BDF(2) scheme displays a different behavior, as seen in Figure 3.12. According to the discrete-dispersion relation, the scheme has considerably more dispersion than the trapezoidal rule. Furthermore, the damping of the scheme results in a smaller final amplitude, even for the long-wave data. On the short scales, the damping is so high that the oscillations have been almost damped out after 5 steps (not shown). At the final time $t_{\mathrm{end}} = 3$ the simulation started with "multiscale" data (Figure 3.12, right) is indistinguishable from the one started with pure long-wave data (Figure 3.12, left). Thus, the scheme is able to balance the short-wave modes that are not resolved in time, but at the same time it suffers from simultaneously damping and dispersing the large-scale data.

The results of the simulations using the multilevel schemes with multiscale initial data are displayed in Figure 3.13. As stated above, the purpose of these schemes is to combine the capability of the trapezoidal rule to resolve the long-wave gravity waves relatively well with a scheme which filters the highly oscillatory short-wave data in an appropriate fashion. The left plot shows the outcome of the method combining the trapezoidal rule with the BDF(2) scheme, whereas the right plot shows the result of the scheme using the trapezoidal rule together with the super-implicit method. For comparison, the result produced by the trapezoidal rule applied only to the long-wave initial data is also shown in the plots (dashed line). As one can see, the latter is nearly identical to the results using the multilevel schemes. This is the desired behavior.
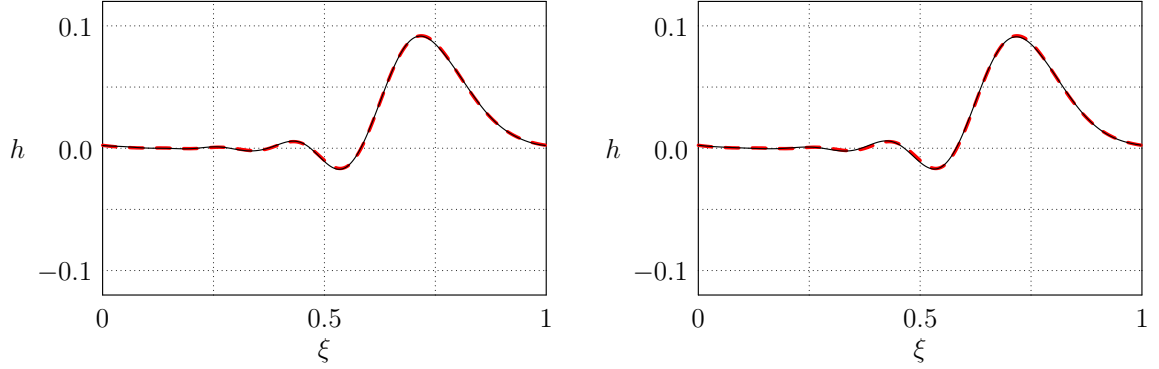
Figure 3.13: Numerical solution (height) of the simple wave test case using the multilevel schemes on a grid with 512 cells, $\mathrm{cfl} = 10$ and "multiscale" initial data at time $t_{\mathrm{end}} = 3$ (black line). Left: results obtained with the multilevel trapezoidal/BDF(2) scheme; right: results obtained with the multilevel trapezoidal/super-implicit scheme. For comparison, the result of trapezoidal rule obtained with only long-wave initial data is plotted as dashed line.

The short-wave data is damped out in such a way that only the long-wave data is left after some time. On the other hand, for a second-order method which involves the solution of only one linear system, this method performs a very accurate integration of the long-wave data.

Finally, it should be mentioned that for the multilevel method with trapezoidal/ BDF(2) base schemes the prolongation operator from (3.16) has been used, whereas in the multilevel trapezoidal/super-implicit scheme (3.17) has been applied. Using (3.16) in the latter one leads to instabilities for reasons unknown so far (not shown).

3.3.2 Balanced modes in presence of time dependent bottom topography

The second test case evaluates the ability of the new multilevel schemes to relax to non-trivial balanced states after some perturbation in presence of time dependent bottom topography. For this purpose, a bottom topography of the form

$$b(t, x) = -\frac{1}{\omega} \cos\left(\omega\, t\right) \tilde{q}(x - x_0)$$

is introduced, where

$$\tilde{q}(x) = \left(\frac{2 + \lambda^2 \sigma^2}{\lambda^2 \sigma^2} \sin(\lambda\,x) + \frac{4x}{\lambda \sigma^2} \cos(\lambda\,x) - \frac{4x^2}{\lambda^2 \sigma^4} \sin(\lambda\,x) \right) \exp\left(-\left(\frac{x}{\sigma}\right)^2 \right) . \quad (3.19)$$

This results in a source term of the form $b_t(t, x) = \sin(\omega\,t)\,\tilde{q}(x - x_0)$. The parameters are given by $\omega = 0.2\pi$, $x_0 = 50$, $\sigma = 10$ and $\lambda = 0.32\pi$. Furthermore, the initial conditions are set to $h(0, x) \equiv u(0, x) \equiv 0$. Here, the focus is on the dynamics on the small scale x. Thus, to obtain a well resolved source term in space, the computations are done on the domain $x \in [0, 100]$ with periodic boundary conditions and a grid with 256 cells (i.e., $\Delta x = 100/256$). In the computations, the source term is integrated analytically over one time step.

As a result of the asymptotic analysis in Section 2.2, height should scale with Fr^2 and velocity should be of order 1 in this case, and as a result of equation (2.14) they should, up to higher order perturbations, have the form

$$h_{\mathrm{asy}}(t, x) = \frac{\mathsf{Fr}^2}{H_0} \omega \cos(\omega\,t)\,\tilde{h}(x - x_0) \qquad (3.20)$$

with

$$\tilde{h}(x) = \frac{1}{\lambda^2} \sin(\lambda\,x) \exp\left(-\left(\frac{x}{\sigma}\right)^2 \right) , \qquad (3.21)$$

and

$$u_{\mathrm{asy}}(t, x) = \frac{1}{H_0} \sin(\omega\,t)\,\tilde{u}(x - x_0) , \qquad (3.22)$$

where

$$\tilde{u}(x) = \left(-\frac{1}{\lambda} \cos(\lambda\,x) + \frac{2x}{\sigma^2 \lambda^2} \sin(\lambda\,x) \right) \exp\left(-\left(\frac{x}{\sigma}\right)^2 \right) . \qquad (3.23)$$

From (3.20) and (3.22) it can be seen that for $t = 0$ only the zero initial condition in u is in agreement with the asymptotic solution. The zero initial condition for the height introduces deviations of order Fr^2 from the balanced state. The Courant number is specified to scale with $1/\mathsf{Fr}$, so that on a grid associated to the large scale ξ the gravity wave speed would always be of order 1. For the computations presented below the background height and the Froude number are set to $H_0 = 1$ and $\mathsf{Fr} = 0.01$ if not stated otherwise. The Courant number is given by $\mathrm{cfl} = 0.8/\mathsf{Fr} = 80$, and the blending factors for the multilevel schemes are $\mu_\nu = (2/3, 1/2, 1/3, 1/6, 0)$.

In Figure 3.14 the results of the simulation using the trapezoidal rule after the first six time steps are displayed. For comparison, the asymptotic solution is plotted as a dashed line. Clearly, the numerical solution does not relax to the asymptotic solution in the height variable. It rather oscillates around the balanced state. This is also true for later time steps (not shown). The results suggest that the balanced velocity field is well approximated. However, as stated above, the initial velocity field is set to be equal to the asymptotic solution. Since the deviation from the balanced solution is very small for the velocity, the non-vanishing deviations are barely visible. This issue will be further analyzed below. The results are in good agreement with our findings from Section 3.1.4.

The outcome of the simulations using the BDF(2) scheme are given in Figure 3.15. In this case the balanced state is essentially attained already after one time step. After three time steps, the numerical solution is nearly indistinguishable from the asymptotic solution. This is the expected behavior for this test case.

Figures 3.16 and 3.17 display the results obtained with the multilevel schemes. For both versions the correct behavior is also observed after a few time steps. In comparison with the reference schemes the solutions after one time step are somewhere between the ones of the trapezoidal rule and the BDF(2) scheme. However, after about five time steps the solutions have relaxed to the asymptotic solution. Again, for the multilevel trapezoidal/BDF(2) method the prolongation operator from (3.16) has been chosen, and for the multilevel trapezoidal/super-implicit scheme the one from (3.17), although in this case the results are hardly distinguishable.

To show that not only the numerical solution of the height variable can exhibit large deviations from the balanced state, the simulations are repeated, but this time they start at $t = \frac{\pi}{4\omega}$. At this time the balanced state is different from the initial data in both variables. The results are shown again after the first six time steps. As shown in Figure 3.18 for the trapezoidal rule there is a much bigger difference in the height variable, which additionally grows during the first time steps (note the different scaling in h compared to the other figures!). This time an additional non-vanishing deviation in the velocity variable is visible, which oscillates around the balanced state. Also the results obtained with the BDF(2) scheme (Figure 3.19) show a considerably larger deviation from the balanced state at the beginning of the simulation. However, after six time steps the solution has again almost relaxed to the balanced state.

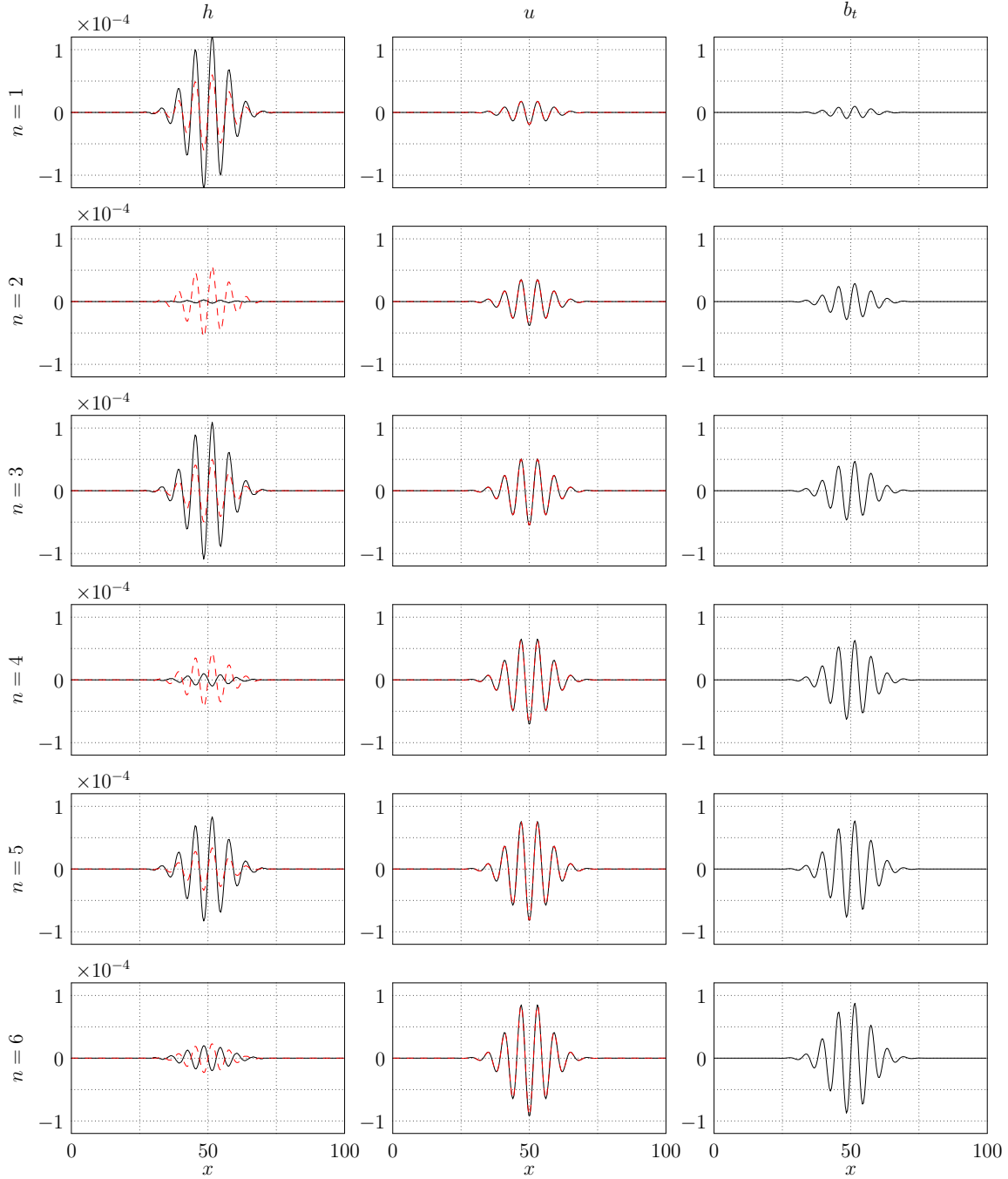**Figure 3.14:** Numerical solution of the balancing test case after the first six time steps using the trapezoidal rule (black line) on a grid with 256 cells, Fr = 0.01 and cfl = 80. Left column: height, middle column: velocity, right column: source term. Each step n is one row. Asymptotic solution is plotted as dashed line.
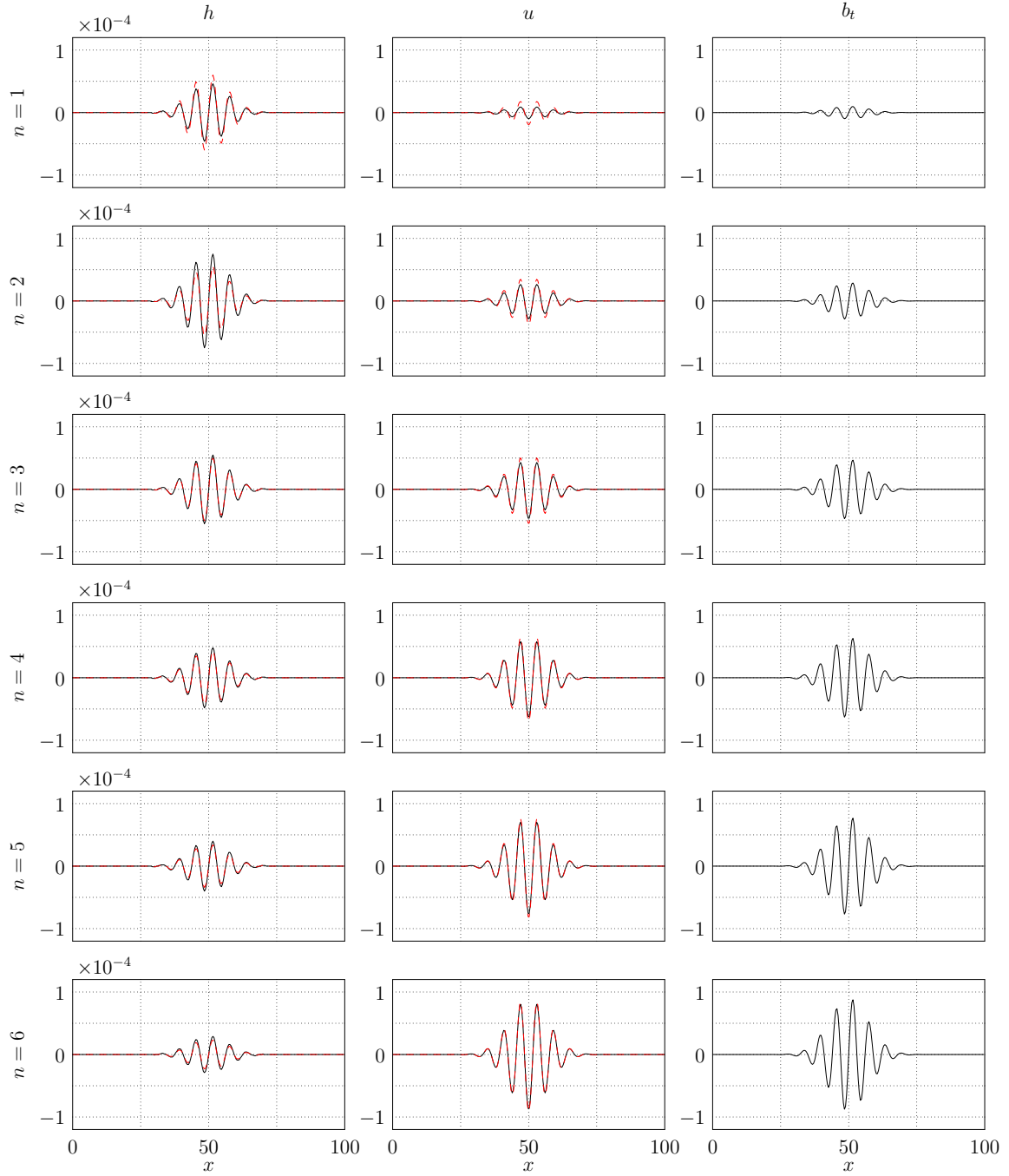
Figure 3.15: Same as Figure 3.14, but using the BDF(2) scheme.

This behavior is also obtained with the multilevel schemes, as displayed in Figures 3.20 and 3.21. For the trapezoidal/BDF(2) scheme the relaxation process takes also about six steps. Additionally, some small artifacts arising from the splitting can be observed during the relaxation process (e.g., the solution after steps three and four). The results from the multilevel trapezoidal/super-implicit scheme show that the relaxation process takes some more steps. Here, the numerical solution is in balance after about 11 steps (not shown). Also for this method some artifacts due to the splitting are visible. Finally, the results for this method using the prolongation operator from (3.16) instead of (3.17) are given in Figure 3.22 after the first three time steps and after step 47 to 49. In this case, a relaxation of the height variable to the balanced state is hardly visible during the first steps. By looking at the later steps, it seems that the scheme excites long-wave modes, which oscillate around the balanced state and become larger in time. A possible reason for this could be that the super-implicit scheme is unstable for long-wave data, and, due to the scale-splitting, "energy" is transferred from the short-wave modes to the long-wave modes. However, due to the good results using the other prolongation operator, this issue was not further investigated.

Note that the results for the multilevel schemes differ from the ones presented in Vater *et al.* (2011). This is due to a wrong calculation of the blending weight μ_ν in the cited reference.

Figure 3.16: Same as Figure 3.14, but using the multilevel trapezoidal/BDF(2) scheme.
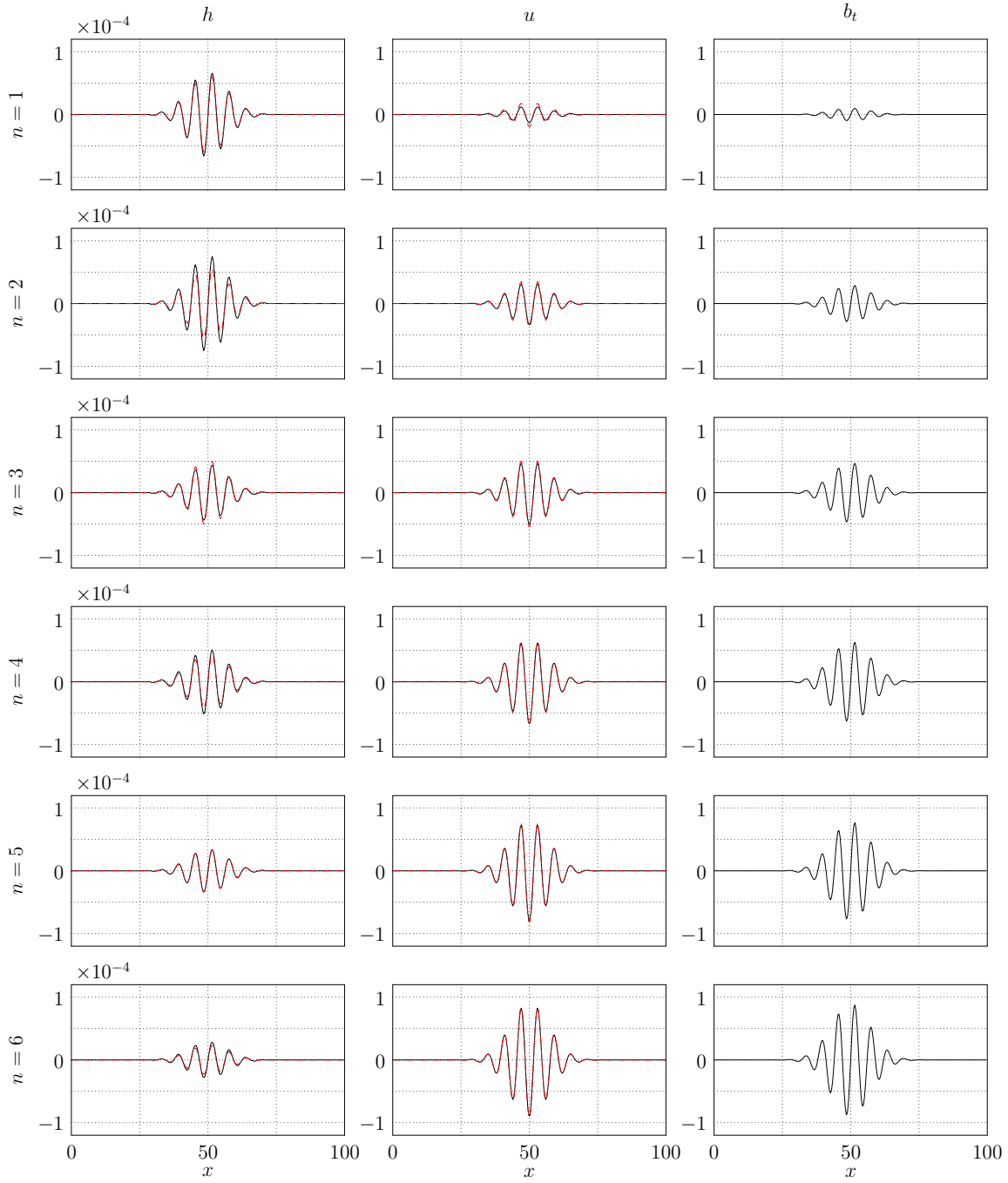
Figure 3.17: Same as Figure 3.14, but using the multilevel trapezoidal/super-implicit scheme.

Figure 3.18: Numerical solution of the balancing test case with completely unbalanced initial data after the first six time steps using the trapezoidal rule (black line) on a grid with 256 cells, $\mathrm{Fr} = 0.01$ and $\mathrm{cfl} = 80$. Notice the different scaling in h compared to the other figures. Asymptotic solution is plotted as dashed line.
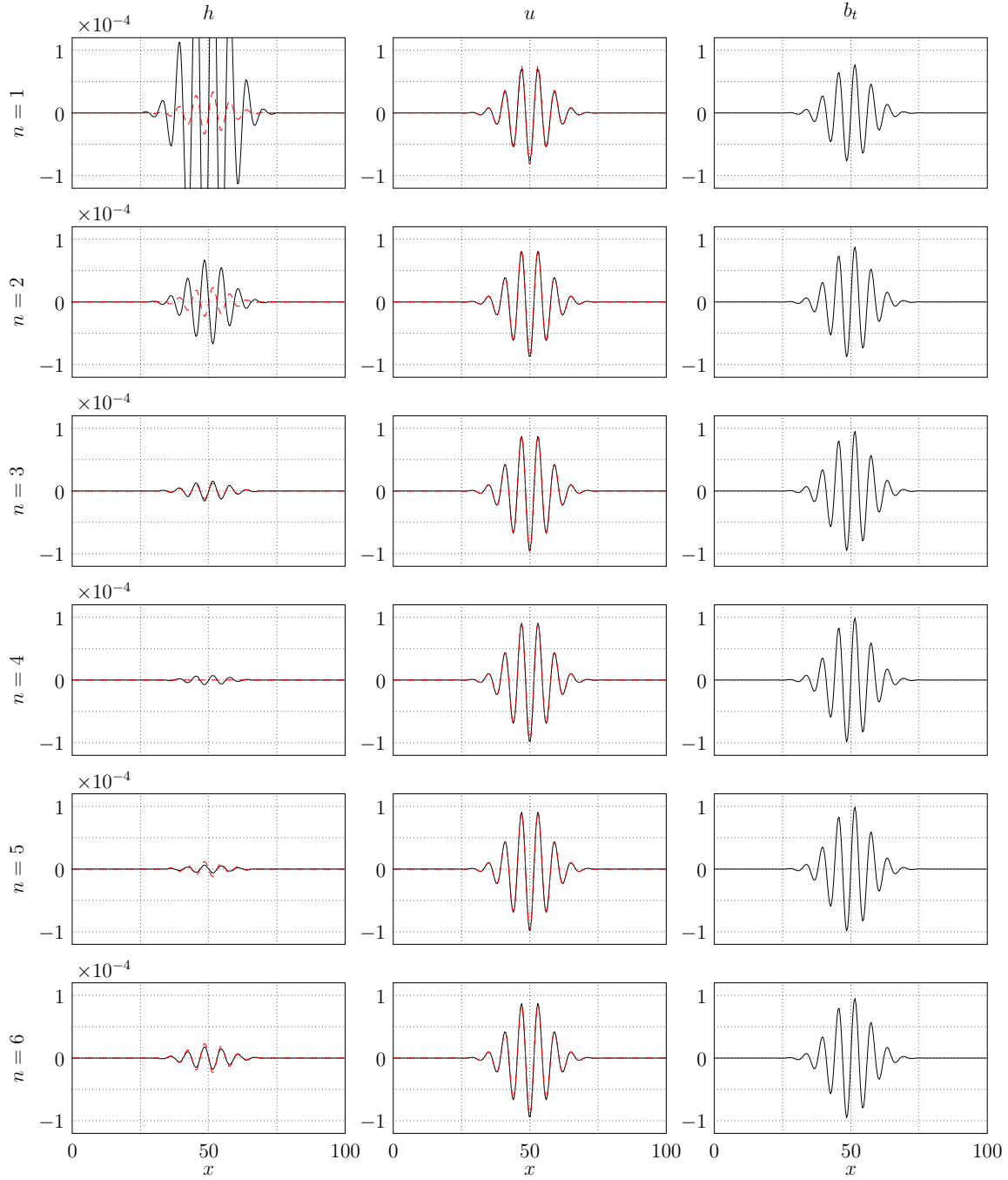
Figure 3.19: Same as Figure 3.18, but using the BDF(2) scheme.
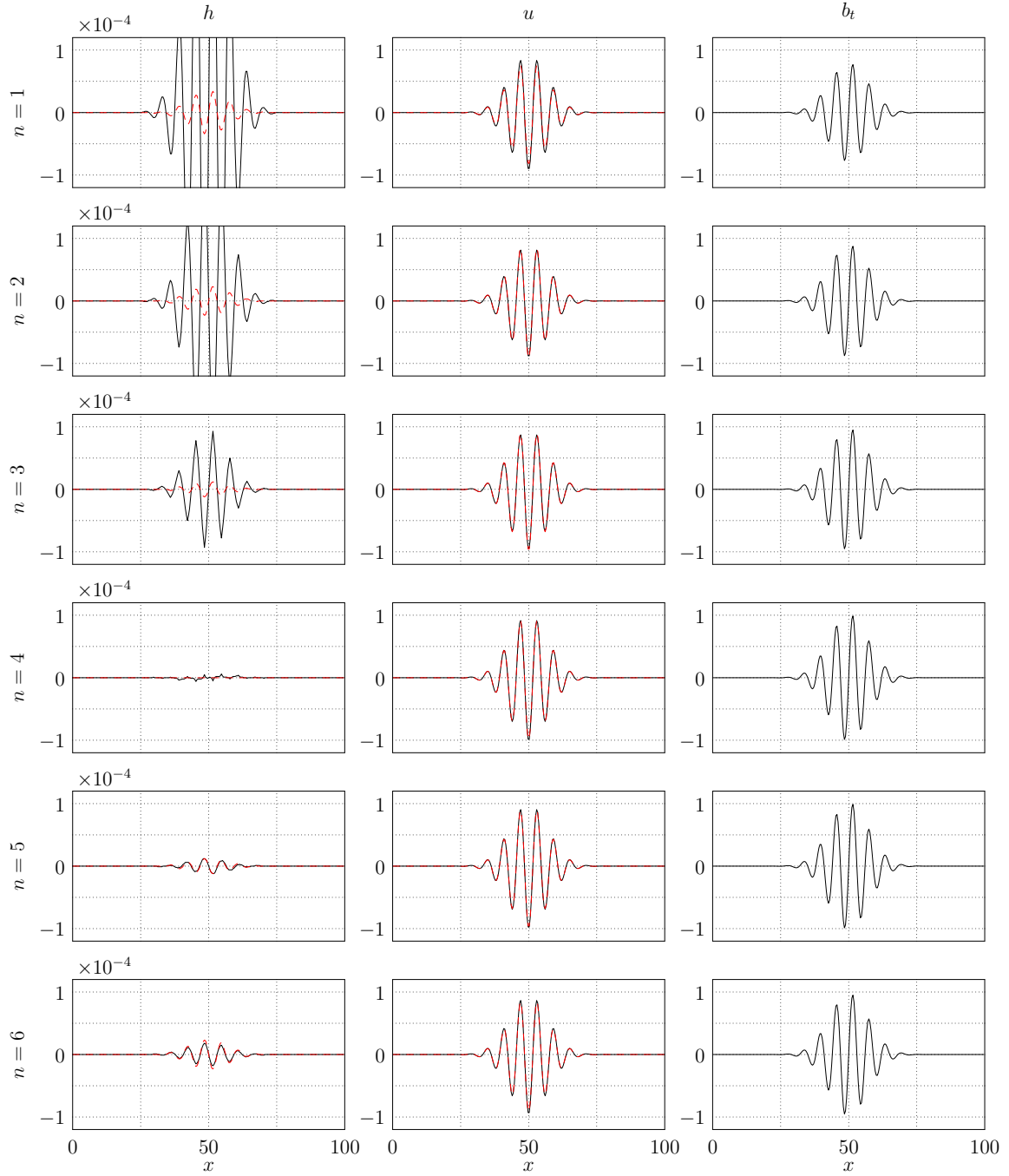
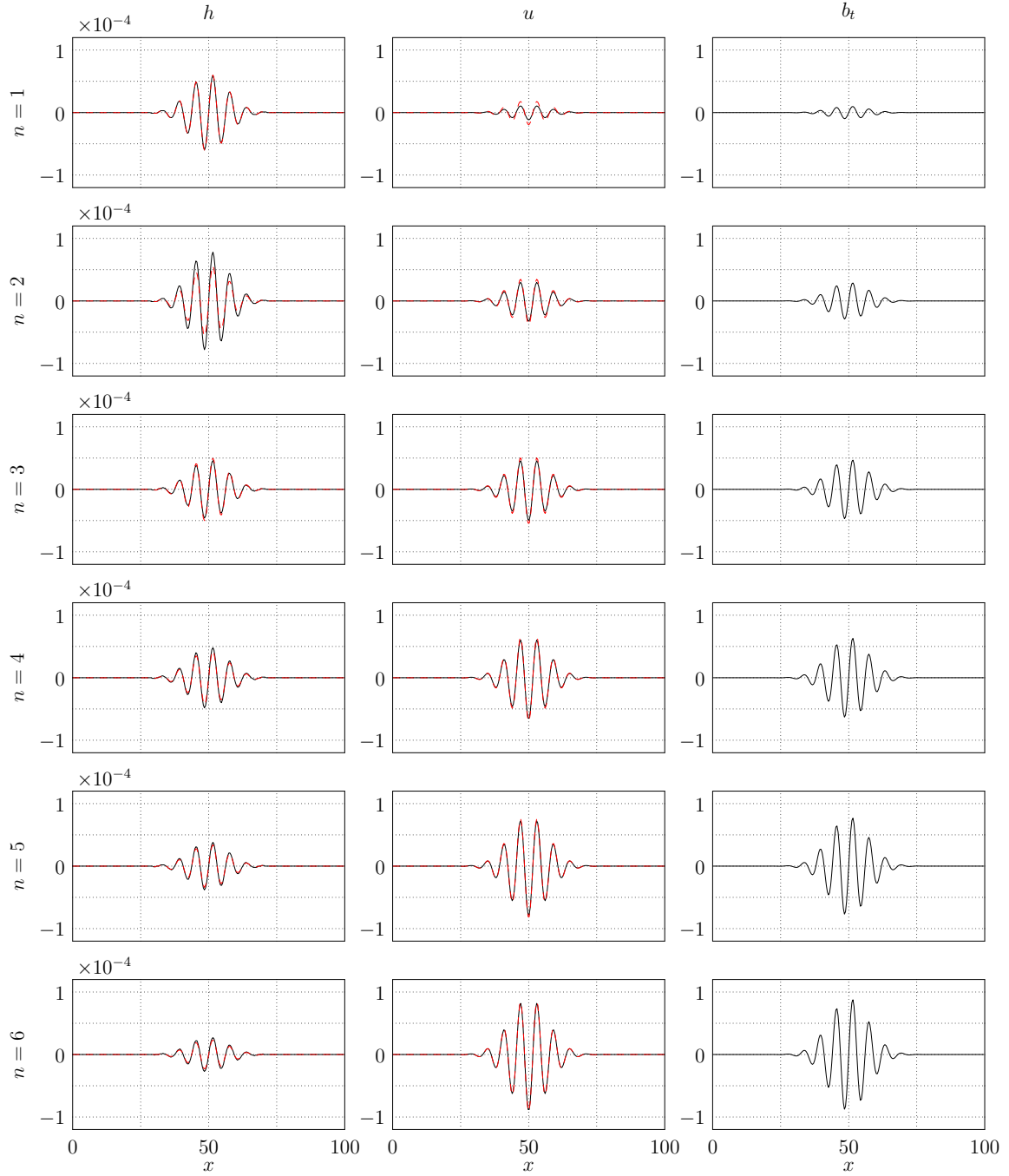**Figure 3.20:** Same as Figure 3.18, but using the multilevel trapezoidal/BDF(2) scheme.

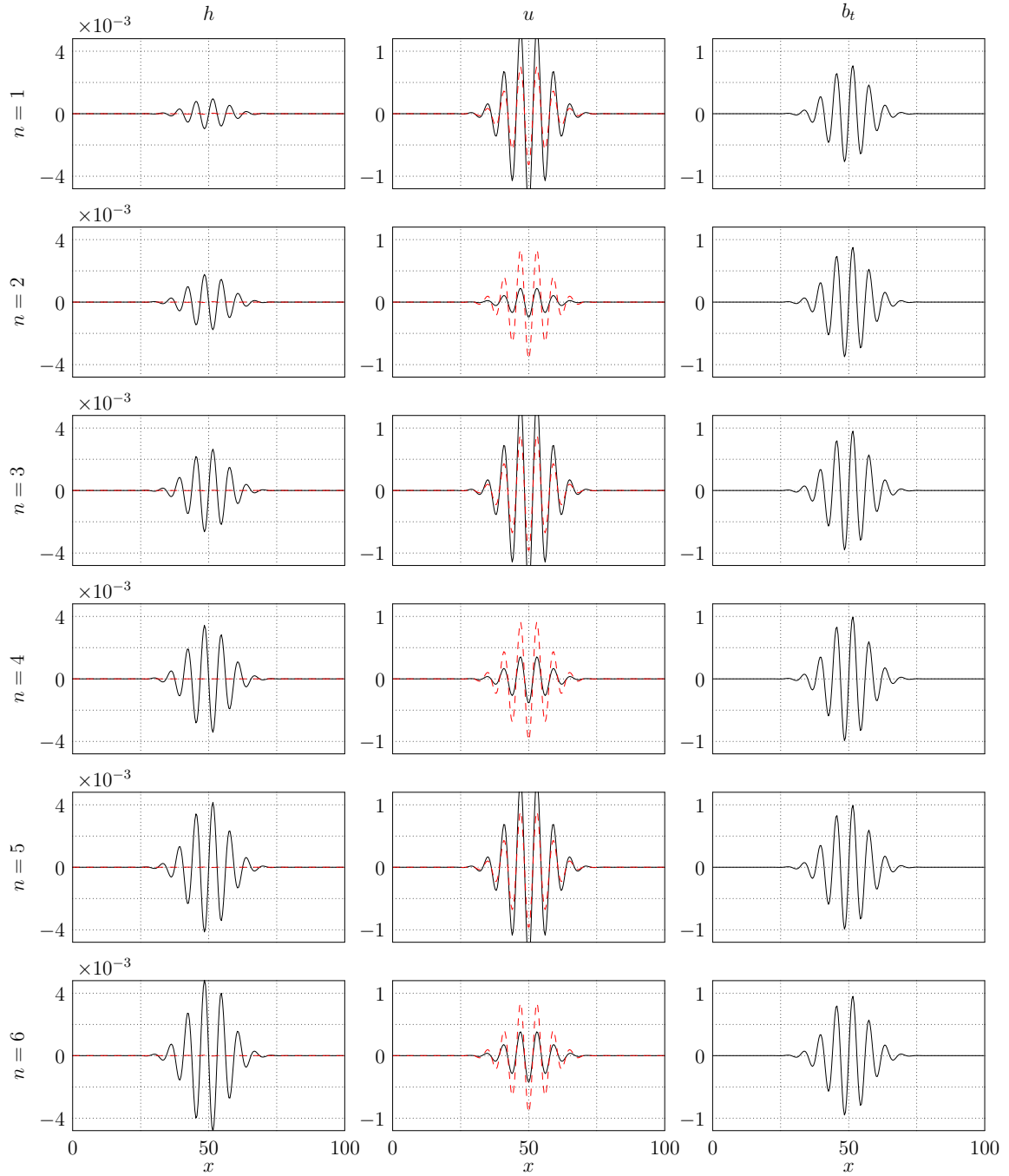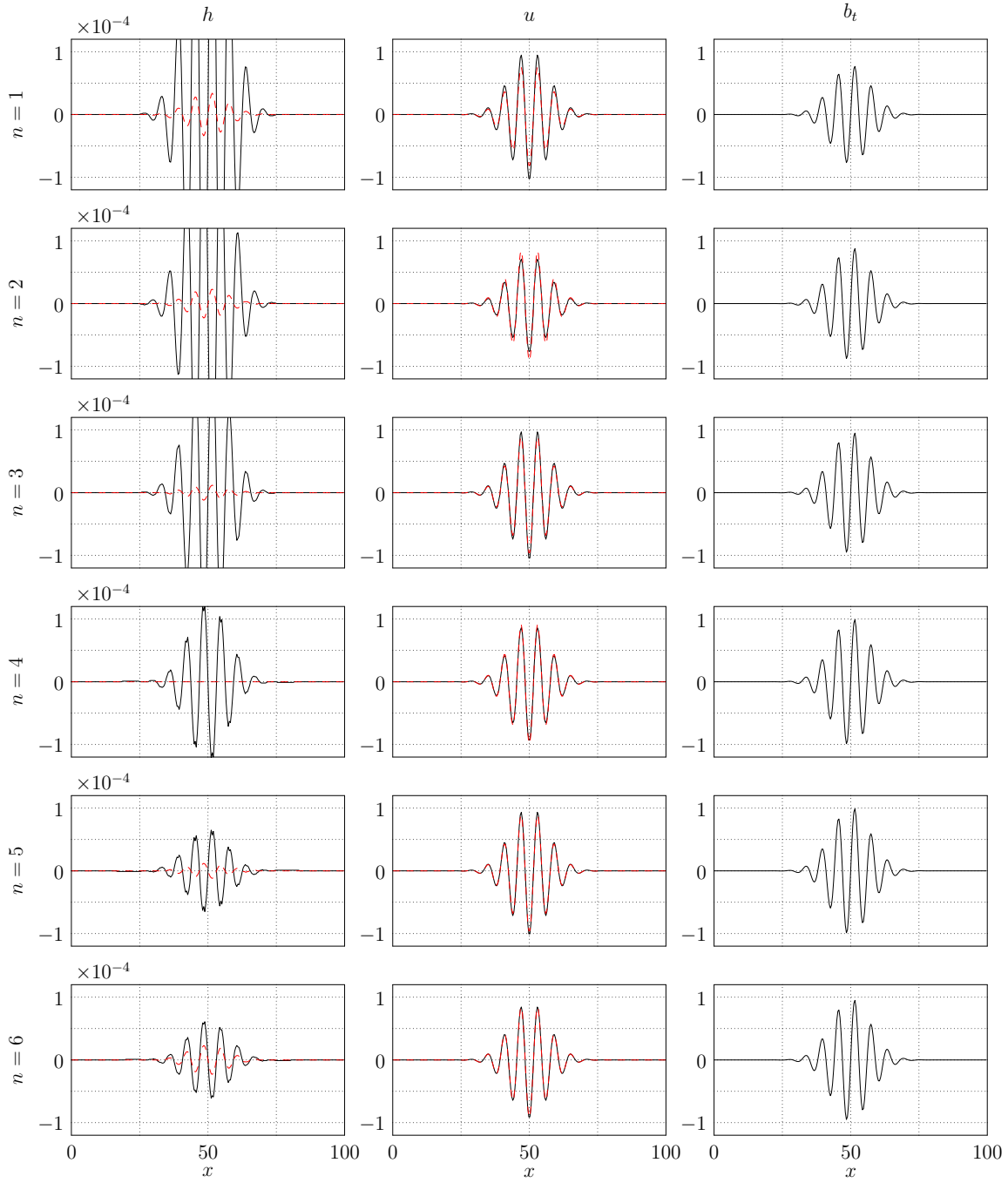**Figure 3.21:** Same as Figure 3.18, but using the multilevel trapezoidal/super-implicit scheme with the prolongation operator (3.17).
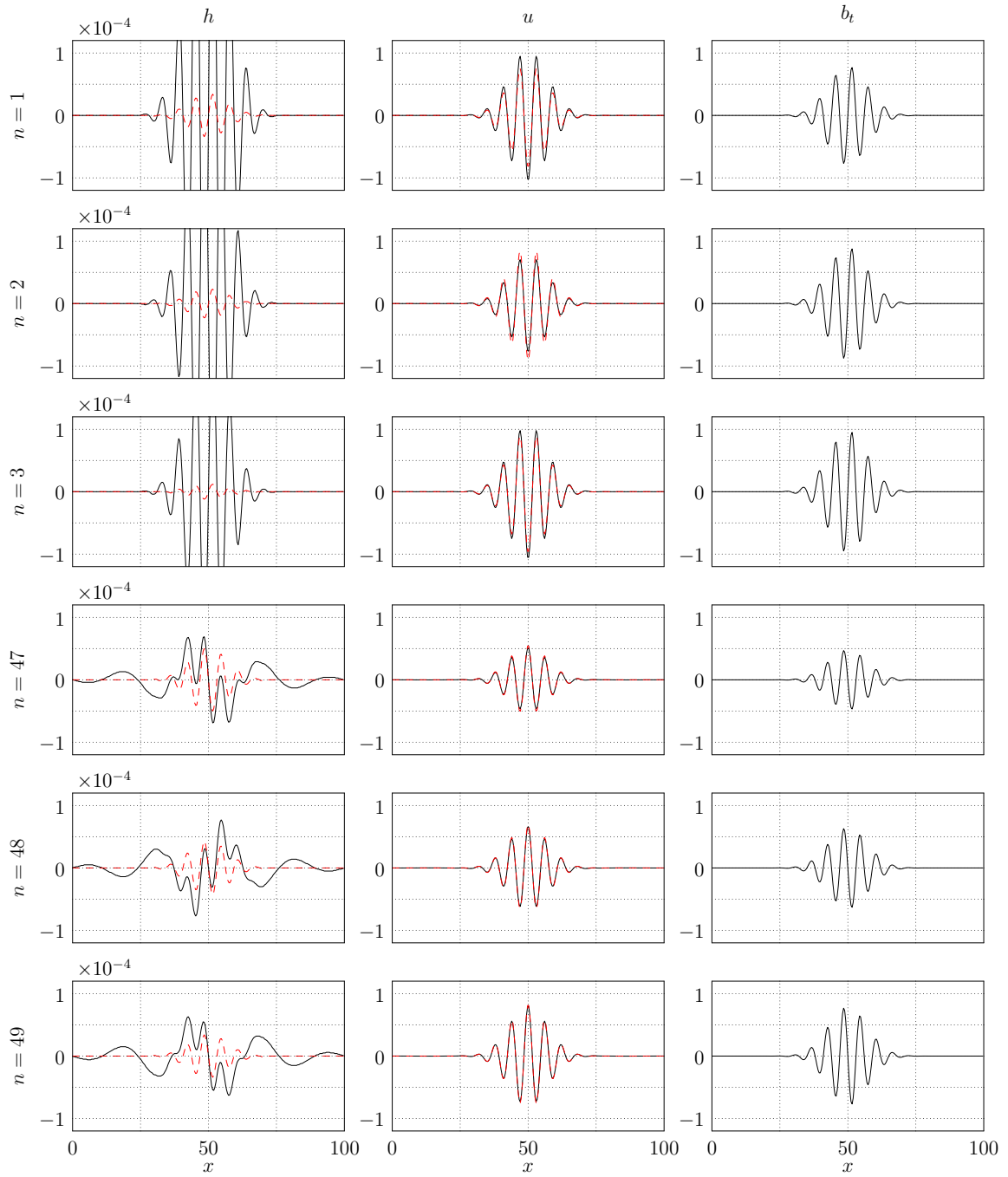
Figure 3.22: Same as Figure 3.18 after the time steps 1, 2, 3, 47, 48, 49, but using the multilevel trapezoidal/super-implicit scheme with the prolongation operator (3.16).

4 Numerical solution of the "lake equations"

In this chapter a projection method for the zero Froude number shallow water equations is constructed, which incorporates bottom topography varying in time. As stated in Section 2.3, this system is also known as the lake equations. It is the second building block for the derivation of the semi-implicit multiscale method for low Froude number shallow water flows developed in Chapter 5. The presented scheme is based on previous works for zero Froude number shallow water flows (Vater, 2005; Vater and Klein, 2009) and for the zero Mach number Euler equations and variants (Schneider *et al.*, 1999; Klein, 2009). It consists of three steps: In a first predictor step an auxiliary system is solved explicitly using standard Godunov-type methods for hyperbolic conservation laws. A MAC-type ("Marker-And-Cell") projection is applied to the advective fluxes in a second step. This ensures that the velocity terms arising in the fluxes are in compliance with the divergence constraint. In a final second projection the non-convective part of the momentum flux is computed in order to have new time level cell-centered velocities which satisfy the divergence constraint, as well.

The development of projection methods for both, viscous and inviscid incompressible flows has a long tradition. It started with the fundamental work of Chorin (1968) and Temam (1968), and the topic is still under high development (see van Kan, 1986; Bell *et al.*, 1989; Bell and Marcus, 1992; Schneider *et al.*, 1999; Almgren *et al.*, 2000; Guermond *et al.*, 2006; Kadioglu *et al.*, 2008, and references therein).

4.1 Derivation of the projection method

Starting from the governing equations, which were discussed in Section 2.3 and are given by

$$h_t + \nabla \cdot (h\boldsymbol{u}) = 0$$
$$(h\boldsymbol{u})_t + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u}) + h\nabla h^{(2)} = 0 \qquad (4.1)$$
$$h(t, \boldsymbol{x}) = H(t) - b(t, \boldsymbol{x})$$

the auxiliary system which is employed in the predictor step is defined by

$$h_t + \nabla \cdot (h\boldsymbol{u}) = 0$$
$$(h\boldsymbol{u})_t + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u}) + \Phi h\nabla(h + b) = -(1 - \Phi)(h\nabla h^{(2)})^{\text{old}} \ ,$$

where the parameter $\Phi \in [0, 1]$ has been introduced. For $\Phi = 1$, the standard shallow water equations with Froude number equal to 1 are recovered, while for $\Phi = 0$ the lake equations are obtained with the two exceptions that the height is a prognostic variable (computed through the first equation) and the pressure gradient $(h\nabla h^{(2)})^{\text{old}}$ is obtained as "source term", which is independent of time (over one time step) and computed from an old (known) time level. The remaining homogeneous system to solve are the pressureless equations introduced in Section 2.5.

The case $\Phi = 1$ has essentially been used in Schneider *et al.* (1999), Vater (2005) and Vater and Klein (2009). It has the advantage that the system to solve in the predictor step is hyperbolic, for which well developed numerical methods exist. On the other hand, fake gravity waves are introduced by the pressure gradient term, which influence the size of the maximum allowed time step. Furthermore, a well-balancing for the source term arising from the bottom topography should be considered. In the case $\Phi = 0$, the fake gravity waves do not exist, and the time step of the explicit method is essentially restricted by the advective terms. Also, no source term is required to be balanced, and, as we will see, a second-order accurate approximation of the momentum is obtained. The only "issue" that remains is the solution of the degenerated system. This choice of Φ was suggested in Klein (2009).

In the following, only the case $\Phi = 0$ is considered, i.e., the auxiliary system has the form

$$h_t + \nabla \cdot (h\boldsymbol{u}) = 0 \ ,$$
$$(h\boldsymbol{u})_t + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u}) = -(h\nabla h^{(2)})^{\text{old}} \ . \qquad (4.2)$$

To distinguish the solution of the governing equations from the solution of the auxiliary system, the latter is labeled with an asterisk from now on. Then, starting with the same initial conditions for the systems (4.1) and (4.2) and setting $(h\nabla h^{(2)})^{\mathrm{old}}(\boldsymbol{x}) = (h\nabla h^{(2)})(t_0, \boldsymbol{x})$, the difference in tendency of the momentum of the two systems is

$$\frac{\partial}{\partial t}[(h\boldsymbol{u}) - (h\boldsymbol{u})^*](x, t_0) = 0 \ . \tag{4.3}$$

This means that solving the auxiliary system with identical initial data at time t_0 leads to deviations of order Δt^2 at time $t_0 + \Delta t$ compared to the solution of the lake equations.

Note, that in relation to auxiliary projection methods, as described in Kadioglu *et al.* (2008), system (4.2) forms the so-called *auxiliary variable equations.* As in the cited reference, the old time pressure gradient is used in the momentum equation as "source term". However, as described below, here a Runge-Kutta time discretization is used instead of a spectral deferred correction method. Additionally, Kadioglu *et al.* do not employ a first projection for the correction of the advective fluxes.

For the derivation of the numerical scheme let us consider a semi-discretization in time from time level t^n to $t^{n+1} \coloneqq t^n + \Delta t$. To simplify the notation $h(\boldsymbol{x}, t^n)$ is abbreviated by h^n etc. A second-order accurate approximation is then obtained by evaluating the flux terms at the half time levels $t^{n+1/2} \coloneqq t^n + \Delta t/2$, i.e.,

$$h^{n+1} = h^n - \Delta t\Big[\nabla \cdot (h\boldsymbol{u})^{n+1/2}\Big] \tag{4.4}$$

and

$$(h\boldsymbol{u})^{n+1} = (h\boldsymbol{u})^n - \Delta t\Big[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2} + (h\nabla h^{(2)})^{n+1/2}\Big] \ . \tag{4.5}$$

Let us suppose that the solution of the auxiliary system with the same initial conditions at time level t^n as for the lake equations is known over this time step. Because of (4.3) momentum and velocity of the auxiliary system at the half time level are second-order accurate approximations. However, they do not fulfill the divergence constraint. Thus, a height correction $\delta h_{\mathrm{fl}}^{(2),n}$ is introduced to correct the flux terms, which is of order Δt, and momentum and velocity are then approximated by

$$(h\boldsymbol{u})^{n+1/2} = (h\boldsymbol{u})^{*,n+1/2} - \frac{\Delta t}{2}h^n\nabla\delta h_{\mathrm{fl}}^{(2),n} \tag{4.6}$$

and

$$\boldsymbol{u}^{n+1/2} = \boldsymbol{u}^{*,n+1/2} - \frac{\Delta t}{2} \nabla \delta h_{\mathrm{fl}}^{(2),n} \ . \tag{4.7}$$

In order to obtain an equation for $\delta h_{\mathrm{fl}}^{(2),n}$, the divergence is applied to (4.6), and the resulting equation is combined with the height equation from (4.1). This results in a first Poisson-type equation for the height correction. It is given by

$$\begin{aligned} \frac{\Delta t}{2} \nabla \cdot (h^n \nabla \delta h_{\mathrm{fl}}^{(2),n}) &= h_t^{n+1/2} + \nabla \cdot (h\boldsymbol{u})^{*,n+1/2} \\ &= H_t^{n+1/2} - b_t^{n+1/2} - h_t^{*,n+1/2} \ , \end{aligned} \tag{4.8}$$

where the height equation of the auxiliary system was used to substitute the momentum divergence by the time derivative of the height. Note, that the right hand side of this equation represents the divergence error at the half time step, which is introduced by the auxiliary system. The term $h_t^{n+1/2}$ on the right hand side is computed through $H_t^{n+1/2}$ and $b_t^{n+1/2}$, the former being determined by the integral flux through the boundary of the entire domain and the mean evolution of the bottom topography (cf. equation (2.17)). Having the solution of (4.8), the new time level height and the first flux term in the momentum update (4.5) can be computed using (4.6) and (4.7). However, a second-order approximation of the non-convective part of the momentum flux is still missing. This is obtained by enforcing the divergence constraint on the new time level momentum field. Let

$$(h\boldsymbol{u})^{**} := (h\boldsymbol{u})^n - \Delta t \left[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2} + h^n \nabla h^{(2),n} + \frac{\delta h^n}{2} \nabla h^{(2),n} \right] \tag{4.9}$$

be an intermediate momentum update, where $\delta h^n := h^{n+1} - h^n$. Then, the momentum at the new time level is computed by

$$(h\boldsymbol{u})^{n+1} = (h\boldsymbol{u})^{**} - \frac{\Delta t}{2} h^{n+1/2} \nabla \delta h^{(2),n} \ , \tag{4.10}$$

where $\delta h^{(2),n} := h^{(2),n+1} - h^{(2),n}$ is the update of the second-order height. The divergence constraint is formulated by a trapezoidal time discretization of the height

equation from system (4.1), which is

$$\frac{1}{2}\left[\nabla \cdot (h\boldsymbol{u})^{n+1} + \nabla \cdot (h\boldsymbol{u})^n\right] = -h_t^{n+1/2} \ . \tag{4.11}$$

Combining this equation with (4.10) finally leads to a Poisson-type equation for $\delta h^{(2),n}$. It has the form

$$\frac{\Delta t}{2}\,\nabla \cdot (h^{n+1/2}\nabla \delta h^{(2),n}) = \nabla \cdot (h\boldsymbol{u})^{**} + \nabla \cdot (h\boldsymbol{u})^n + 2\,H_t^{n+1/2} - 2\,b_t^{n+1/2} \ . \tag{4.12}$$

With the solution of this problem, the non-advective contribution to the momentum update can be fully computed. The new momentum is calculated through (4.10), and it satisfies the divergence constraint in the form (4.11). In a last step, the height perturbation at the new time level $h^{(2),n+1}$ is computed by

$$h^{(2),n+1} = h^{(2),n} + \delta h^{(2),n} \ .$$

Apart from the spatial discretization, this procedure describes the projection method for zero Froude number shallow water flows. Compared to Vater (2005) and Vater and Klein (2009), the main differences are the different auxiliary systems solved in the predictor step, and the fact that the height h is no longer constant in the Poisson-type equations. This results in a weighted Laplace operator, as it is usually the case in projection methods for the zero Mach number Euler equations with variable density.

It should be noted that the new time level height could be simply computed by the third equation of (4.1) in each time step. This is not intended at this point, since the ultimate goal is to extend the described method to non-zero Froude numbers, as done in Chapter 5. In this case, the height must be computed explicitly, and doing it here as well, avoids the introduction of errors with respect to this implementation detail at later times.

The scheme is solved on a Cartesian mesh with grid cells $V_{i,j}$. Furthermore, a dual discretization is introduced, where each dual grid cell $\bar{V}_{i+1/2,j+1/2} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$ is centered around a node $(x_{i+1/2}, y_{j+1/2})$ of the primary grid (see Figure 4.1). The
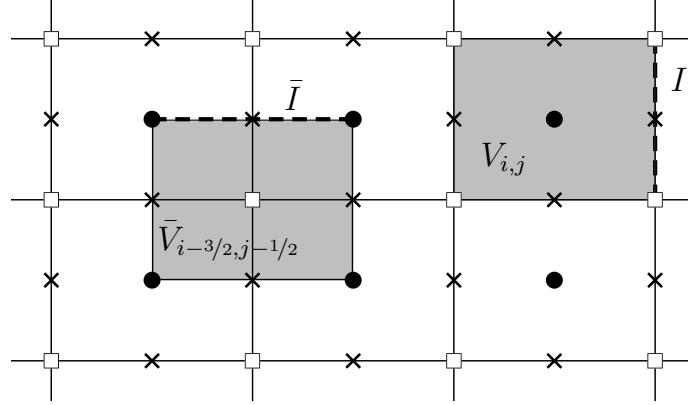
Figure 4.1: Control volumes and interfaces of the primary and dual discretizations. Cell centers of the primary grid are denoted by *circles*, nodes by *squares* and midpoints of the interfaces by *crosses*.

whole method is discretized as a finite volume method, which has the general form

$$\mathbf{U}_{i,j}^{n+1} = \mathbf{U}_{i,j}^{n} - \frac{\Delta t}{|V_{i,j}|} \sum_{I \in \mathcal{I}_{\partial V_{i,j}}} |I| \, \mathbf{F}_{I}^{n+1/2} + \Delta t \, \mathbf{S}_{i,j}^{n+1/2} \; . \tag{4.13}$$

Here, $|V_{i,j}|$ is the volume of cell $V_{i,j}$, and $\mathcal{I}_{\partial V_{i,j}}$ is the set of interfaces being part of the boundary of this cell. $\mathbf{U}_{i,j}^{n}$ represents an approximation to the cell mean of the unknowns $(h, h\boldsymbol{u})$ in the cell $V_{i,j}$ at time t^n, and $\mathbf{F}_{I}^{n+1/2}$ is the numerical flux across the interface I. The latter approximates the average of the flux function

$$\mathbf{f}(\mathbf{u}(t, \boldsymbol{x}), \boldsymbol{n}(\boldsymbol{x})) = \begin{pmatrix} h(\boldsymbol{u} \cdot \boldsymbol{n}) \\ h\boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{n}) + H \, h^{(2)} \, \boldsymbol{n} \end{pmatrix} \tag{4.14}$$

over one time step $[t^n, t^{n+1}]$ up to second-order accuracy. The additional source term $\mathbf{S}_{i,j}^{n+1/2}$ arises in the presence of bottom topography and is an approximation to $\mathbf{s}(\mathbf{u}(t, \boldsymbol{x}), b(t, \boldsymbol{x})) = (0, b\nabla h^{(2)})^T$. The equations are discretized to obtain a scheme which is in conservation form for the height equation. For the exact equations, conservation of momentum is only valid when no bottom topography is present. In this case, also momentum should be conserved on the discrete level. Following the above (semi-discrete) derivation of the scheme, numerical fluxes and the source term

are computed in three steps:

$$\mathbf{F}_I^{n+1/2} := \mathbf{F}_I^{*,n+1/2} + \mathbf{F}_I^{\mathsf{MAC}} + \mathbf{F}_I^{\mathsf{P2}} \ ,$$
$$\mathbf{S}_{i,j}^{n+1/2} := \mathbf{S}_{i,j}^{*,n+1/2} + \mathbf{S}_{i,j}^{\mathsf{MAC}} + \mathbf{S}_{i,j}^{\mathsf{P2}} \ . \tag{4.15}$$

In this formulation, $\mathbf{F}_I^*$ and $\mathbf{S}_{i,j}^{*,n+1/2}$ are the numerical flux and source term approximating the flux function and source term of the auxiliary system, respectively. These are

$$\mathbf{f}^* = \begin{pmatrix} h(\boldsymbol{u} \cdot \boldsymbol{n}) \\ h\boldsymbol{u}(\boldsymbol{u} \cdot \boldsymbol{n}) + (Hh^{(2)})^{\mathrm{old}}\boldsymbol{n} \end{pmatrix} \quad \text{and} \quad \mathbf{s}^* = \begin{pmatrix} 0 \\ (b\nabla h^{(2)})^{\mathrm{old}} \end{pmatrix} \ . \tag{4.16}$$

The second flux term

$$\mathbf{F}_I^{\mathsf{MAC}} := -\frac{\Delta t}{2} \begin{pmatrix} h^n \nabla \delta h_{\mathrm{fl}}^{(2),n} \cdot \boldsymbol{n} \\ (h\boldsymbol{u})^{*,n+1/2}\nabla \delta h_{\mathrm{fl}}^{(2),n} \cdot \boldsymbol{n} + h^n \nabla \delta h_{\mathrm{fl}}^{(2),n}\boldsymbol{u}^{*,n+1/2} \cdot \boldsymbol{n} \end{pmatrix}_I + \begin{pmatrix} 0 \\ \frac{\delta H^n}{2} h^{(2),n}\boldsymbol{n} \end{pmatrix}_I \tag{4.17}$$

corresponds to the first correction computed by equation (4.8). As stated above, with this correction the new time level height can be determined, which enables us to calculate a first contribution to the non-convective part of the flux (represented by the term within the second parentheses in (4.17)) and the source term $\mathbf{S}_{i,j}^{\mathsf{MAC}} = (0, \frac{\delta b^n}{2}\nabla h^{(2),n})^T$. The third terms in of (4.15) are given by

$$\mathbf{F}_I^{\mathsf{P2}} := \begin{pmatrix} 0 \\ H^{n+1/2}\frac{\delta h^{(2),n}}{2}\boldsymbol{n} \end{pmatrix}_I \quad \text{and} \quad \mathbf{S}_{i,j}^{\mathsf{P2}} := \begin{pmatrix} 0 \\ b^{n+1/2}\nabla \frac{\delta h^{(2),n}}{2} \end{pmatrix}_{i,j}$$

and represent the correction computed by the second Poisson-type equation (4.12). Note, that in the practical implementation of the scheme, flux and source term might not be strictly separated, although it must be ensured that there is no contribution from the part representing the source term in the absence of bottom topography.

For the discretization of the bottom topography b, a piecewise bilinear distribution on each primary grid cell which is continuous across the interfaces is assumed. Therefore, b is defined by its values at the nodes of the primary grid, and the cell mean value on $V_{i,j}$ is equal to b at the point (x_i, y_j). Furthermore, the time derivatives $b_t^{n+1/2}$ are approximated by the finite difference quotient composed with the values at

the full time levels:

$$\frac{\partial}{\partial t} b_{i+1/2,j+1/2}^{n+1/2} \approx \frac{b_{i+1/2,j+1/2}^{n+1} - b_{i+1/2,j+1/2}^{n}}{\Delta t} \qquad \text{and} \qquad \frac{\partial}{\partial t} b_{i,j}^{n+1/2} \approx \frac{b_{i,j}^{n+1} - b_{i,j}^{n}}{\Delta t} \ .$$

4.1.1 Solution of the auxiliary system

In the predictor step the auxiliary system (4.2) is solved using a Godunov-type method for hyperbolic conservation laws (van Leer, 1979). As mentioned above, these are the pressureless equations with the "source term" $(h^n \nabla h^{(2),n})$ in the momentum equation. Note, that this term involves not only the contributions from the bottom topography, but also the non-convective part of the flux in (4.16). For the integration, a semi-discretization in space with second-order reconstruction in the primitive variables and Runge-Kutta time stepping is used (Osher, 1985). In particular, Heun's method is applied to obtain second-order accuracy in time. This method is often chosen, since it is *total variation diminishing* (TVD), or, more generally, *strong stability preserving* (SSP) (Shu and Osher, 1988; Gottlieb *et al.*, 2001). The numerical fluxes are evaluated by solving the exact Riemann problem of the pressureless equations at the cell interfaces.

Remark 4.1 *The exact fluxes for the one-dimensional Riemann problem applied to the pressureless equations*

$$(h_0(x), u_0(x)) = \begin{cases} (h_j, u_j) & \text{for } x < x_{j+1/2} \\ (h_{j+1}, u_{j+1}) & \text{for } x \geq x_{j+1/2} \end{cases}$$

are given by

$$F_{i+1/2} = \begin{cases} (h_j u_j, h_j u_j^2) & \text{if } u_\delta > 0 \\ (h_{j+1} u_{j+1}, h_{j+1} u_{j+1}^2) & \text{if } u_\delta < 0 \\ ((h_j u_j + h_{j+1} u_{j+1})/2, h_j u_j^2 = h_{j+1} u_{j+1}^2) & \text{if } u_\delta = 0 \ . \end{cases} \qquad (4.18)$$

for $u_j \geq u_{j+1}$ (Bouchut et al.*, 2003). Here, u_δ is the Roe-average of the velocity,*

defined in (2.23). *For the case* $u_j \leq u_{j+1}$, *the exact fluxes have the form*

$$
F_{i+1/2} = \begin{cases} (h_j u_j, h_j u_j^2) & \textit{if } u_j > 0 \\ (h_{j+1} u_{j+1}, h_{j+1} u_{j+1}^2) & \textit{if } u_{j+1} < 0 \\ (0, 0) & \textit{otherwise .} \end{cases} \tag{4.19}
$$

It is easy to verify that the first two cases in (4.18) *and* (4.19), *correspond to the HLLE flux (Einfeldt, 1988) of the shallow water equations in the limiting case* $\mathsf{Fr} \rightarrow \infty$. *This could be a useful property in the further development of an all Froude number shallow water scheme, where the scheme switches to an explicit Godunov-type method for sufficiently large Froude numbers.* ◁

The source term of the auxiliary system takes h^n and $h^{(2),n}$ as input variables at the beginning of the time step. The spatial discretization should be the same as the one for the second correction. This will be described in the next section.

4.1.2 Spatial discretization of the correction equations

In the first correction, the flux divergence of the auxiliary system is essentially corrected by the gradient of $\delta h_{\mathrm{fl}}^{(2),n}$, which results in a MAC-type projection (Harlow and Welch, 1965; Welch *et al.*, 1965). To be consistent, the same standard finite volume flux divergence is applied in the Poisson-type equation (4.8) for the discretization of the divergence, as it has been used in the flux computation of the auxiliary system (see also Figure 4.2, left). It is given by

$$
\mathsf{D}_{i,j}(\boldsymbol{u}) := \frac{1}{|V_{i,j}|} \int_{\partial V_{i,j}} \boldsymbol{u} \cdot \boldsymbol{n} \; \mathrm{d}\sigma = \frac{1}{|V_{i,j}|} \sum_{I \in \mathcal{I}_{\partial V_{i,j}}} \int_I \boldsymbol{u} \cdot \boldsymbol{n} \; \mathrm{d}\sigma \; . \tag{4.20}
$$

The discretization of $\delta h_{\mathrm{fl}}^{(2),n}$ is the same as in Süli (1991) and Vater (2005), which assume a piecewise bilinear distribution of this quantity on the dual grid. The gradient of such a function is piecewise linear with discontinuities along the interfaces of the dual grid cell. However, on the interfaces of the primary grid, where the integral arising from the divergence must be computed, this gradient is well defined. The height h^n is assumed to be piecewise constant on each interface I. The value is
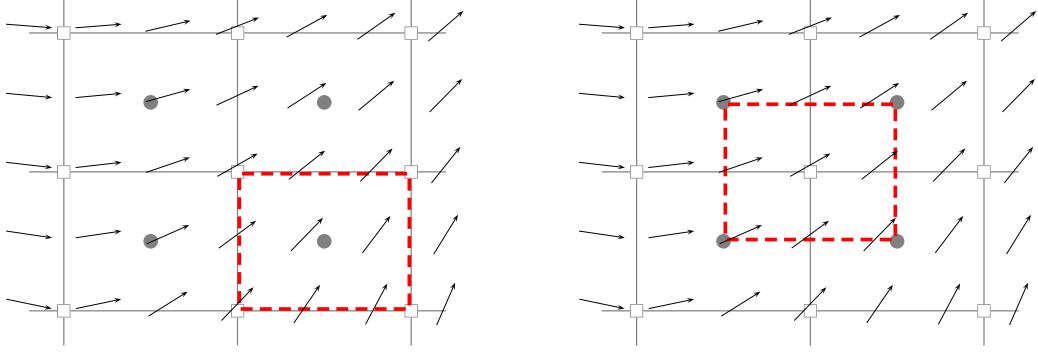
Figure 4.2: Application of the divergence constraint in the MAC (left) and the second projection (right).

interpolated by hydrostatic reconstruction from the two neighboring cells. This means that we use the fact that $H(t) = h(t, \boldsymbol{x}) + b(t, \boldsymbol{x})$ should be constant in space. Then, it follows for the interface between the cells $V_{i,j}$ and $V_{i+1,j}$ that

$$h_{i+1/2,j}^n = h_{i,j}^n + b_{i,j}^n - b_{i+1/2,j}^n = h_{i+1,j}^n + b_{i+1,j}^n - b_{i+1/2,j}^n \ .$$

To account for numerical errors, these values are reconstructed from both sides of the interface and averaged.

Remark 4.2 *Other choices are possible. In Klein (2009) simply an average of the two adjacent cell averages of the analogous values in the case of an anelastic model was taken. Numerical experiments show that there is not much difference in the results. However, the hydrostatic reconstruction seemed to be the most plausible approach in the case of the zero Froude number shallow water equations.* ◁

Let us denote the hydrostatically reconstructed value of h on an interface I by h_I. Then, the discretized Poisson-type problem (4.8) can be written in the form

$$\frac{\Delta t}{2} \frac{1}{|V_{i,j}|} \sum_{I \in \mathcal{I}_{\partial V_{i,j}}} h_I^n \int_I \nabla \delta h_{\mathrm{fl}}^{(2),n} \cdot \boldsymbol{n} \, \mathrm{d}\sigma = \frac{H^{n+1} - H^n}{\Delta t} - \frac{b_{i,j}^{n+1} - b_{i,j}^n}{\Delta t} - \frac{h_{i,j}^{*,n+1} - h_{i,j}^{*,n}}{\Delta t} \quad (4.21)$$

for all i and j, where the time derivatives on the right hand side of (4.8) have been approximated by the differences between the values of the full time levels. Since the

degrees of freedom for the solution $\delta h_{\text{fl}}^{(2),n}$ are located in the center of each cell $V_{i,j}$, one linear equation is obtained for each degree.

For the computation of the flux correction as in (4.17) also the momentum and velocity of the auxiliary system at the interfaces is needed. In the current implementation, these are interpolated in space and time from the cell centered values of the neighboring cells at the full time levels with second-order accuracy. As stated in Klein (2009), other approaches are possible. In particular, the author showed that upwinding has the advantage of obeying the maximum principle for advected scalars. However, since the computations in the present work do not involve the advection of additional scalars, this implementation detail is of minor importance.

The correction of the non-convective terms in the fluxes and the source term is done separately by adding $-\Delta t(\frac{\delta h^n}{2}\nabla h^{(2),n})$ to the momentum equation. The spatial discretization has again to be the same as it is described below for the second projection.

Note, that the terms $\delta h_{\text{fl}}^{(2),n}$ and $h^{(2),n}$ appearing in the first and second parentheses of (4.17), respectively, are two unrelated numerical quantities. The former quantity is the correction computed in the first projection. The latter quantity is the full height perturbation, which is subsequently updated by the outcome of the second projection, and has been already used in the computation of the "source term" for the auxiliary system.

For the second correction, the divergence constraint is applied to each dual control volume, as can be seen in Figure 4.2, right. This leads to a divergence defined by

$$
\begin{aligned}
\bar{\mathsf{D}}_{i+1/2,j+1/2}(\boldsymbol{u}) &:= \frac{1}{|\bar{V}_{i+1/2,j+1/2}|}\int_{\partial\bar{V}_{i+1/2,j+1/2}} \boldsymbol{u}\cdot\boldsymbol{n}\ \mathrm{d}\sigma \\
&= \frac{1}{|\bar{V}_{i+1/2,j+1/2}|}\sum_{\bar{I}\in\bar{\mathcal{I}}_{\partial\bar{V}_{i+1/2,j+1/2}}}\int_{\bar{I}}\boldsymbol{u}\cdot\boldsymbol{n}\ \mathrm{d}\sigma\ .
\end{aligned}
\tag{4.22}
$$

Also the correction $\delta h^{(2),n}$ is assumed to be piecewise bilinear, but this time on the primary control volumes. Therefore, the gradient of such a quantity is piecewise linear on each cell. Moreover, it needs to be defined how the height which enters as weight in the Laplacian of the left hand side of (4.12) is discretized. Since the mean values of the height at the full time levels are already known before the second projection,

$h^{n+1/2} = (h^n + h^{n+1})/2$ is simply computed by linear interpolation in time. For the spatial discretization it is assumed that the height is piecewise constant on each cell. This leads to a piecewise linear distribution of $h^{n+1/2} \nabla \delta h^{(2),n}$ on each grid cell, and the integral terms to compute the weighted Laplacian with help of the divergence (4.22) are well defined.

As in Vater (2005) and Vater and Klein (2009) the presented projection method is implemented as an exact projection method, meaning that the velocity field at each time step satisfies a discretization of the divergence constraint up to the accuracy of the linear system solver used in the second correction. Following these works, the linear reconstruction of the unknown variables is computed before the second projection (and not in the predictor step) in the present case. This leads to piecewise linear momentum terms $(h\boldsymbol{u})^{**}$ and $(h\boldsymbol{u})^n$ under the divergence on the right hand side of (4.12). The resulting discretized Poisson-type problem is then given by

$$
\frac{\Delta t}{2} \bar{\mathsf{D}}_{i+1/2,j+1/2} (h^{n+1/2} \nabla \delta h^{(2),n}) = \bar{\mathsf{D}}_{i+1/2,j+1/2} \left[(h\boldsymbol{u})^{**} + (h\boldsymbol{u})^n \right]
$$
$$
+ 2 \frac{H^{n+1} - H^n}{\Delta t} - 2 \frac{b^{n+1}_{i+1/2,j+1/2} - b^n_{i+1/2,j+1/2}}{\Delta t} \; . \tag{4.23}
$$

Remark 4.3 *In the absence of bottom topography, a 9-point stencil for the Laplacian is obtained in two space dimensions. This discretization has been shown to be inf-sup stable by Vater and Klein (2009), meaning that any intermediate momentum update before the second projection leads to a unique pressure-momentum pair computed during the second projection. Therefore, no checkerboard instability can occur.* ◁

For the final momentum update, the piecewise linear contributions of $h^{n+1/2} \nabla \delta h^{(2),n}$ on each grid cell are added to the (already reconstructed) intermediate momentum $(h\boldsymbol{u})^{**}$. This completes the description of the projection method for the lake equations. In the next section, some numerical tests illustrate the method's performance in more detail.

4.2 Numerical results

Although the zero Froude number shallow water equations with non-trivial bottom topography are discussed to some extent for practical applications in the literature

(see, e.g., Bühler and Jacobson, 2001; Barreiro and Bühler, 2008; Johnson *et al.*, 2005), not much is known about exact solutions. To derive solutions for time independent bottom topography, it is usually employed that the governing equations satisfy Kelvin's circulation theorem. This implies the material advection of potential vorticity $\omega = b^{-1}\boldsymbol{k} \cdot (\nabla \times \boldsymbol{u})$, i.e.,

$$\omega_t + \boldsymbol{u} \cdot \nabla \omega = 0 \ . \tag{4.24}$$

For stationary solutions, it turns out that stream lines must be parallel to levels of constant vorticity. Therefore, the vorticity can be assumed to be a function of the stream function ψ with $\nabla^\perp \psi = h\boldsymbol{u}$. However, such derived solutions often imply restrictions on the computational domain or the boundary conditions, e.g. the domain is required to be circular. The implementation of these additional features goes beyond the scope of this work.

Test cases with flat bottom topography have been already discussed in Vater (2005). Therefore, only problems involving non-trivial bottom topography are investigated. The considered test cases are the "lake at rest" problem and a quasi one-dimensional test case, where the exact solution is known. In a more sophisticated test case, a vortex is advected over non-trivial topography. Here, the exact solution is unknown, but the results are discussed qualitatively and the experimental order of convergence is computed.

In this work, only periodic boundary conditions are considered. The implementation of wall boundary conditions was discussed in Vater (2005). Since no mass inflow or outflow can occur with periodic boundary conditions the background height is set constant $H \equiv 1$ in all test cases, and the term H_t is set to 0. The Cartesian grid is set to have equal grid spacing $\Delta x = \Delta y$. To obtain an initial momentum field, which satisfies the divergence constraint on a discrete level, i.e.,

$$\bar{\mathsf{D}}_{i+1/2,j+1/2}((h\boldsymbol{u})^0) = b_t(0, \boldsymbol{x}) \ ,$$

first the exact initial solution at the cell centers is linearly reconstructed using central differences to obtain $(h\boldsymbol{u})^{0,r}$. This data is corrected by the solution of the Poisson

problem

$$\bar{\mathsf{D}}_{i+\nicefrac{1}{2},j+\nicefrac{1}{2}}(\nabla\varphi) = \bar{\mathsf{D}}_{i+\nicefrac{1}{2},j+\nicefrac{1}{2}}((h\boldsymbol{u})^{0,r}) - b_t(0,\boldsymbol{x}) \ .$$

The initial data is then given by

$$(h\boldsymbol{u})^0_{i,j} = (h\boldsymbol{u})^{0,r}_{i,j} - \nabla\varphi_{i,j} \ .$$

In the practical computations, the term $b_t(0,\boldsymbol{x})$ is approximated by a centered difference in time with a very small step size of 10^{-8}. Note, that this momentum field is piecewise linear on each grid cell $V_{i,j}$, but discontinuous across the interfaces. The initial height perturbation is set to $h^{(2)}(0,x,y) \equiv 0$, and the correct value is only available after the computation of the first time step through $\delta h^{(2),0}$.

Since the auxiliary system is computed explicitly, the stability of its solution strongly depends on a CFL time step restriction (Courant *et al.*, 1928). In the computations it is ensured that the time step always satisfies $\Delta t \leq \text{cfl}\,\Delta t^{\text{max}}$ with $0 < \text{cfl} < 1$ and

$$\Delta t^{\text{max}} \coloneqq \min_{i,j}\left(\frac{1}{\Delta t^{\text{max}}_{x,i,j}} + \frac{1}{\Delta t^{\text{max}}_{y,i,j}}\right)^{-1} \ ,$$

where $\Delta t^{\text{max}}_{x,i,j} = \min(\Delta t^{\text{max}}_{x-,i,j}, \Delta t^{\text{max}}_{x+,i,j})$ (and similarly $\Delta t^{\text{max}}_{y,i,j}$) is the maximal possible time step for an explicit update of cell $V_{i,j}$ in the x-direction (y-direction) based on the fluxes through the left ($I_{x-,i,j}$) and right ($I_{x+,i,j}$) interface (see Botta *et al.* (2004) for more details). $\Delta t^{\text{max}}_{x-,i,j}$ and $\Delta t^{\text{max}}_{x+,i,j}$ are computed by the rule

$$\Delta t^{\text{max}}_{x\pm,i,j} \coloneqq \frac{|V_{i,j}|}{|I_{x\pm,i,j}||\boldsymbol{u}_{x\pm,i,j}\cdot\boldsymbol{n}|} = \frac{|V_{i,j}|}{|I_{x\pm,i,j}||u_{\delta,x\pm,i,j}|} \ ,$$

where $u_{\delta,x\pm,i,j}$ is the Roe average of the interface velocities and the (only) eigenvalue of the Jacobian of the flux function of the pressureless equations (see equation (2.23)). Note, that on a uniform Cartesian grid with grid spacing Δx, the estimated time step is $\Delta t^{\text{max}} = \mathcal{O}(\Delta x)$. The parameter cfl, which is called *Courant* or *CFL* number, is a safety factor and is set to at most 0.9 in all computations.

The linear systems (4.21) and (4.23) for the solution of $\delta h^{(2),n}_{\text{fl}}$ and $\delta h^{(2),n}$ are solved using a matrix-free implementation of the Bi-CGSTAB algorithm (van der Vorst,

1992). In each iteration, the Euclidean norm of the residual vector is calculated, and the algorithm is terminated when either the absolute value or the value relative to the norm of the initial residuum is less then a given tolerance. In most calculations, this tolerance is set to 10^{-10}.

4.2.1 Lake at rest

The first test case is taken from the literature of so-called well-balanced methods (LeVeque, 1998; Audusse *et al.*, 2004), which possess the property that they maintain some balance of source terms on the discrete level. For the initial conditions a fluid at rest ($\boldsymbol{u} \equiv \boldsymbol{0}$) is considered over non-trivial topography. This state should be maintained up to machine accuracy.

In particular, the computational domain is defined to be $\Omega = [0, 1]^2$ with periodic boundary conditions. The bottom topography is given by

$$b(t, x, y) = \tilde{b}(r) \qquad \text{with} \qquad r = \sqrt{(x - 0.5)^2 + (y - 0.5)^2}$$

and

$$\tilde{b}(r) = \begin{cases} a \cdot \frac{\exp(-0.5/(r_m^2 - r^2))}{\exp(-0.5/r_m^2)} & \text{if } r < r_m, \\ 0 & \text{otherwise.} \end{cases}$$

This defines an isolated "hill" with radius r_m and height a. In the simulation, these parameters are set to $r_m = 0.3$ and $a = 0.2$. Then, the initial height and momentum are given by $h(0, x, y) = 1 - b(0, x, y)$ and $(h\boldsymbol{u})(0, x, y) \equiv (0, 0)$, respectively.

With $h^{(2)}(0, x, y) \equiv 0$ at the beginning, the source term becomes zero for the initial step. Furthermore, since the variables of the auxiliary system are reconstructed in h and $\boldsymbol{u}$ in the predictor step, the velocity used in the computation of the numerical fluxes (4.18) and (4.19) is also identical to zero. This leads to no change of the initial state through the solution of the auxiliary system, and therefore, a zero right hand side in the Poisson-type equation (4.8), which provides the first correction of the numerical fluxes. Since this equation is solved using an iterative method with a zero initial iterate, the correction becomes (exactly!) zero, as well. It is easy to check that the same reasoning also holds for the second Poisson-type equation (4.12). Therefore, the "lake at rest" steady state is exactly preserved.
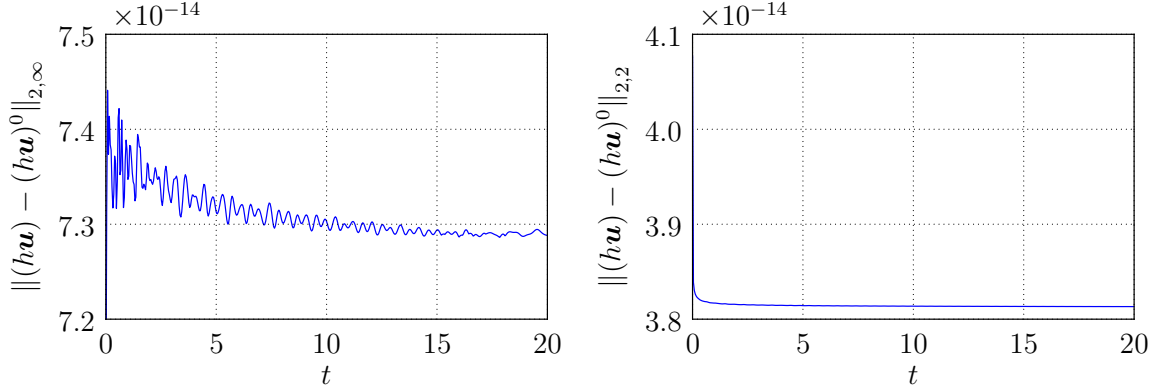
Figure 4.3: Error in momentum for the "lake at rest" test case with initial perturbations of order 10^{-13} measured in the L^∞ and L^2 norms.

This result also holds in numerical experiments, when the initial conditions are in exact balance. However, the method should also keep this balance, when the initial data are slightly perturbed. For this reason, the test case is set up a second time with an initial zero momentum field to which white noise of order 10^{-13} is added. Furthermore, the stopping criterion of the linear system solver is set, such that the solution of the correction equations leads to at least one iteration. The computational grid is chosen to have 256×256 cells, and the time step is $\Delta t = 1/400$. The evolving error over time in the momentum measured in the L^∞ and L^2 norms is displayed in Figure 4.3. The results show that the initial error even reduces a bit until it stabilizes.

4.2.2 Flow over a hill

In a second test case, the flow over a quasi one-dimensional hill is considered. In this case, the exact solution is known, since the divergence constraint becomes trivial, and the momentum is given up to a constant by the time derivative of the height, i.e., for initial data only changing in x,

$$(hu)(0, x, y) = -\int_{x_0}^{x} h_t^0(x)\,\mathrm{d}x + (hu)^0 \quad \text{and} \quad (hv)(0, x, y) = (hv)^0 \,, \qquad (4.25)$$

where h^0 is the initial height distribution in x, $(hu)^0$ the initial x-momentum at x_0, and $(hv)^0$ the initial (constant) momentum in the y-direction. The domain is chosen to be $\Omega = [0, 1]^2$, again with periodic boundary conditions. For the definition of the

bottom topography, let us define the speed v_{rel} at which the topography is advected in the x direction, and set $x_{\text{rel}} \coloneqq x - v_{\text{rel}} \cdot t \pmod 1$, where the term $a \pmod b$ represents the remainder after Euclidean division generalized to real numbers. Then, the topography is given by

$$b(t, x, y) = \tilde{b}(x - v_{\text{rel}} \cdot t \pmod 1) + C = \tilde{b}(x_{\text{rel}}) + C$$

with

$$\tilde{b}(x_{\text{rel}}) = \begin{cases} a \cdot \frac{\exp(-0.5/(r_m^2 - (x_{\text{rel}} - 0.5)^2))}{\exp(-0.5/r_m^2)} & \text{if } |x_{\text{rel}} - 0.5| < r_m, \\ 0 & \text{otherwise.} \end{cases} \tag{4.26}$$

Width and height of the hill are set to $r_m = 0.3$ and $a = 0.2$, respectively. The constant $C = -\overline{\tilde{b}(x_{\text{rel}})}$ is defined such that the mean of the bottom topography is zero. Then, the height of the fluid is given by $h = 1 - b$, and the initial momentum must be computed according to (4.25), which results in

$$(hu)(0, x, y) = (hu)^0 + v_{\text{rel}}(h^0(x) - h^0(x_0)) \quad \text{and} \quad (hv)(0, x, y) = (hv)^0 \,, \tag{4.27}$$

where the quantities to be defined are set to $(hu)^0 - v_{\text{rel}}h^0(x_0) = 1$ and $(hv)^0 = 1$. Changing the value of v_{rel} corresponds to a Galilei transformation between the two associated states.

The numerical solution is computed up to the final time $t_{\text{end}} = 8$ on different computational grids with 32×32, 64×64, 128×128 and 256×256 cells. Furthermore, four different choices of the relative velocity are investigated, in particular $v_{\text{rel}} = -1.0, -0.5, 0.0, +1.0$. For a good comparison between different choices of v_{rel}, the time step on each grid is held constant with $\Delta t_{32} = 1/125$, $\Delta t_{64} = 1/250$, $\Delta t_{128} = 1/500$ and $\Delta t_{256} = 1/1000$, which corresponds to advective Courant numbers $\text{cfl}_{v_{\text{rel}}=-1.0} \approx 0.35$, $\text{cfl}_{v_{\text{rel}}=-0.5} \approx 0.48$, $\text{cfl}_{v_{\text{rel}}=0.0} \approx 0.61$ and $\text{cfl}_{v_{\text{rel}}=1.0} \approx 0.86$.

In Figure 4.4, the distribution of (hu) in x-direction at $y = 0.5$ is shown. The initial data (red dashed line) is plotted together with the approximate solution at $t_{\text{end}} = 8$ (black line) for the 64×64 grid. At the final time, the hill has traveled eight times ($v_{\text{rel}} = -1.0, +1.0$) or four times ($v_{\text{rel}} = -0.5$) through the domain. For $v_{\text{rel}} = 0.0$ the hill is stationary with center $x = 0.5$. In all cases the initial distribution is well preserved. Only for the cases $v_{\text{rel}} = 0.0, +1.0$, a global weak decrease in momentum
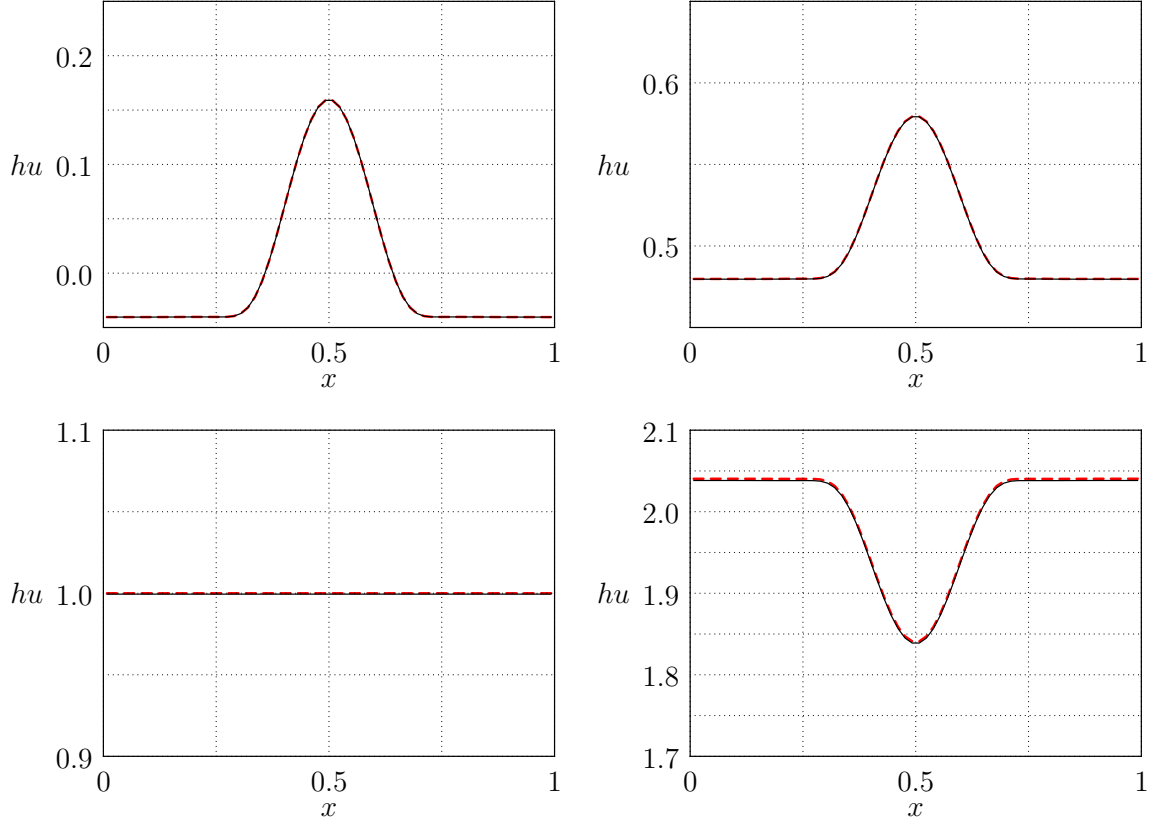
Figure 4.4: Momentum in x-direction for the "flow over a hill" test case at $y = 0.5$ with $v_{\text{rel}} = -1.0$ (top left), $v_{\text{rel}} = -0.5$ (top right), $v_{\text{rel}} = 0.0$ (bottom left), $v_{\text{rel}} = +1.0$ (bottom right). Initial data (red dashed), and momentum at $t = 8$ (black). Grid with 64×64 grid cells.

can be observed. The decrease can be better quantified, when the data is Galilei transformed to the case $v_{\text{rel}} = 0.0$, which is shown in Figure 4.5.

Additionally, the errors and convergence rates are computed for the simulated results. Here, the error vector $\boldsymbol{e}^N$ in the momentum with elements

$$e_{i,j}^N := \sqrt{\left((hu)_{i,j}^0 - (hu)_{i,j}^N\right)^2 + \left((hv)_{i,j}^0 - (hv)_{i,j}^N\right)^2}$$

is calculated by comparing the cell mean values of the initial data with those of the final data. This can be done, since the exact solution is identical for times $t_{\text{ini}} = 0$ and $t_{\text{end}} = 8$. The global error is measured using discrete versions of the L^2 and the
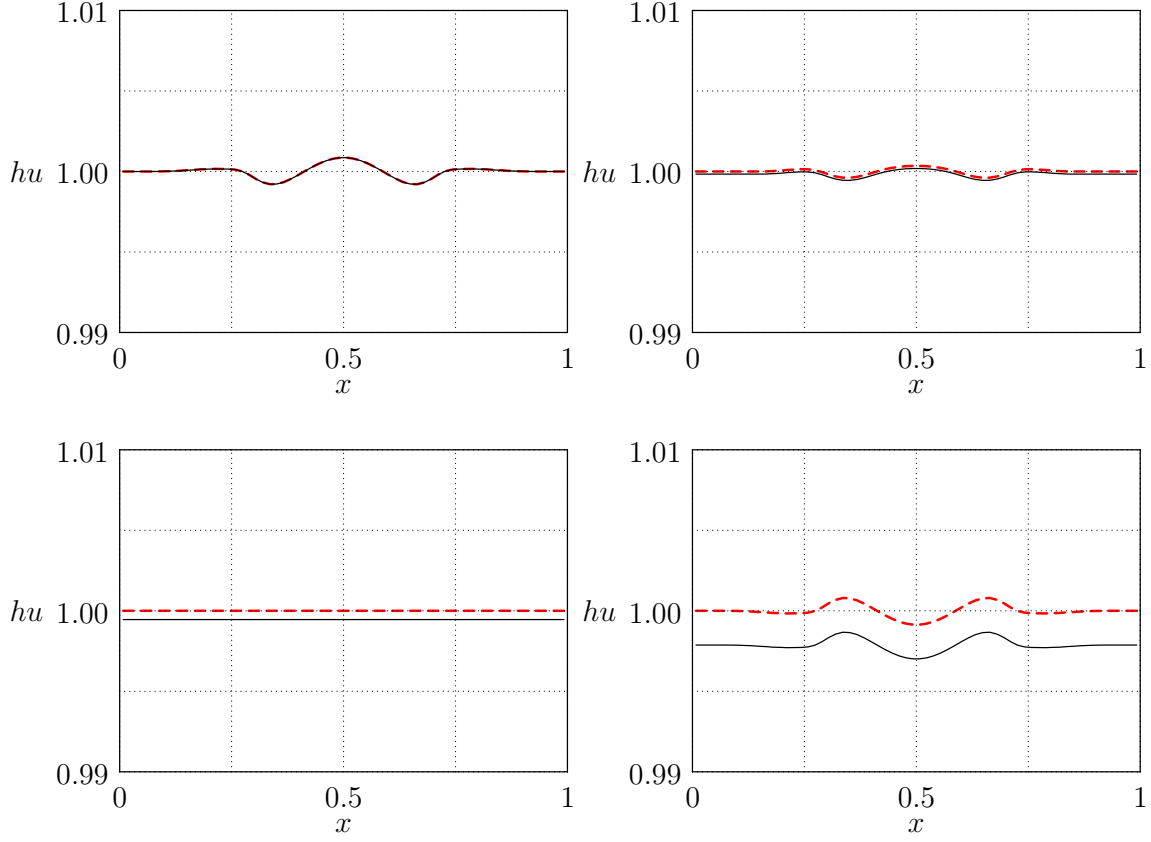
Figure 4.5: Momentum in x-direction for the "flow over a hill" test case at $y = 0.5$, transformed to $v_{\mathrm{rel}} = 0.0$. $v_{\mathrm{rel}} = -1.0$ (top left), $v_{\mathrm{rel}} = -0.5$ (top right), $v_{\mathrm{rel}} = 0.0$ (bottom left), $v_{\mathrm{rel}} = +1.0$ (bottom right). Initial data (red dashed), and momentum at $t = 8$ (black). Grid with 64×64 grid cells.

L^∞ norms. These are defined by

$$\left\| \boldsymbol{e}^N \right\|_2 := \left(\sum_{i,j} |V_{i,j}| \, |e_{i,j}^N|^2 \right)^{1/2} \quad \text{and} \quad \left\| \boldsymbol{e}^N \right\|_\infty := \max_{i,j} \{ e_{i,j}^N \} \ .$$

Additionally, the experimental convergence rate γ is given, which is calculated by the formula

$$\gamma := \frac{\log(\|\boldsymbol{e}_c^N\| / \|\boldsymbol{e}_f^N\|)}{\log(\Delta x_c / \Delta x_f)} \ . \tag{4.28}$$

In this definition, $\boldsymbol{e}_c^N$ and $\boldsymbol{e}_f^N$ are the computed error vectors of the solution on a coarse and a fine grid and Δx_c and Δx_f are the corresponding grid spacings.
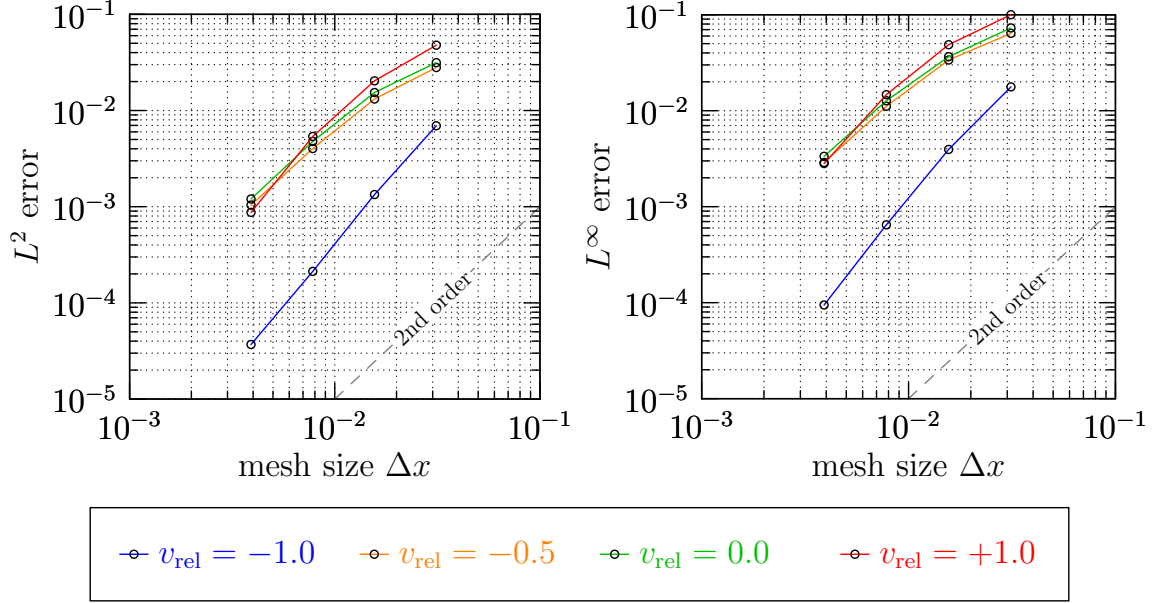
Figure 4.6: Experimental errors in $(h\boldsymbol{u})$ for the "flow over a hill" test case for different choices of v_{rel} at time $t_{\mathrm{end}} = 8$.

The results are shown in Figure 4.6 (see Table A.1 for the precise numerical values). The estimated convergence rates clearly indicate second-order convergence, and for the cases $v_{\mathrm{rel}} = -1.0, +1.0$ they are even a little bit higher on the finest grids. It should be noted that the simplified setup with zero momentum in the y-direction led to super-convergence (not shown). The quasi one-dimensional data probably made the leading terms in the truncation errors of the scheme vanish. This could be also the reason for the higher convergence rates in the case $v_{\mathrm{rel}} = -1.0$. Furthermore, the errors computed for the case $v_{\mathrm{rel}} = -1.0$ are much smaller than for the other test cases, which is probably due to the fact that in this case the mean x-momentum is zero by definition of the initial values in (4.27).

The aforementioned loss of momentum due to numerical diffusion is further investigated by looking at the evolution of the x-momentum in the discrete L^1 norm

$$\left\| (hu)^N \right\|_1 := \sum_{i,j} |V_{i,j}| \, |(hu)^N_{i,j}| \; . \tag{4.29}$$

In Figure 4.7, the relative change of this quantity is plotted for different values of v_{rel}. The analysis shows that the loss in momentum scales with the absolute value of the
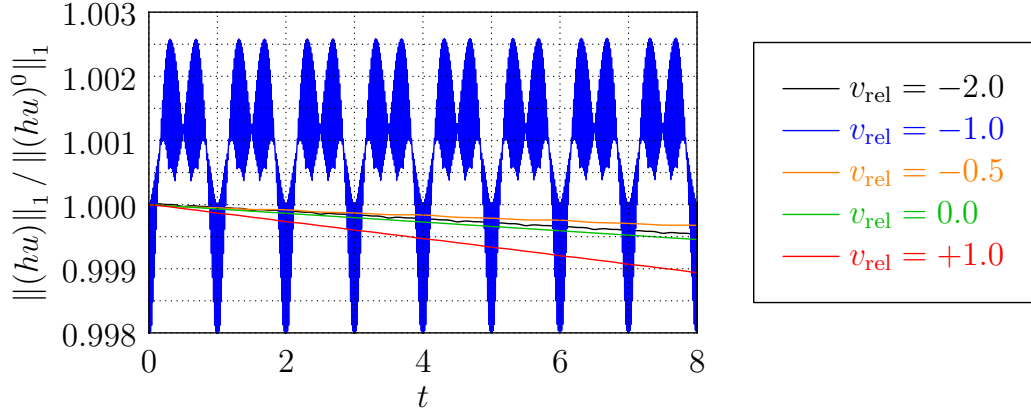
Figure 4.7: Relative change over time of x-momentum in the discrete L^1 norm defined in (4.29) for different choices of v_{rel}. Grid with 64×64 grid cells.

mean momentum. The biggest loss in (hu) can be observed for $v_{\text{rel}} = +1.0$, where the initial mean of (hu) is 2.0. For $v_{\text{rel}} = 0.0$, the initial mean x-momentum is 1.0, and the loss is approximately 50 percent compared to the case $v_{\text{rel}} = +1.0$. For $v_{\text{rel}} = -0.5$ (mean in (hu): $+0.5$), the decrease is about 25 percent. To show that the same holds for cases where the mean is negative, the case with $v_{\text{rel}} = -2.0$ has been added (mean in (hu): -1.0) in Figure 4.7. The loss in momentum is approximately the same as for $v_{\text{rel}} = 0.0$. For the case $v_{\text{rel}} = -1.0$ the initial mean in (hu) is zero, and there is no loss in momentum over time. The value of the L^1 norm merely oscillates around the initial value.

It is also of interest how the loss of momentum scales with grid resolution. In Figure 4.8 the absolute change compared to the initial value of the L^1 norm of (hu) is given for the case $v_{\text{rel}} = -0.5$ on different grids. The results show that a refinement of the grid by a factor 2 leads to a loss in x-momentum by roughly a factor of $8 = 2^3$. This would correspond to third order convergence and is higher than the convergence results in Figure 4.6 (resp. Table A.1).

4.2.3 Vortex advected over topography

In a last test case a vortex, which is advected over non-trivial topography, is considered. For this problem, the exact solution is unknown. However, some qualitative studies are possible at least: Since the lake equations imply the material conservation of potential
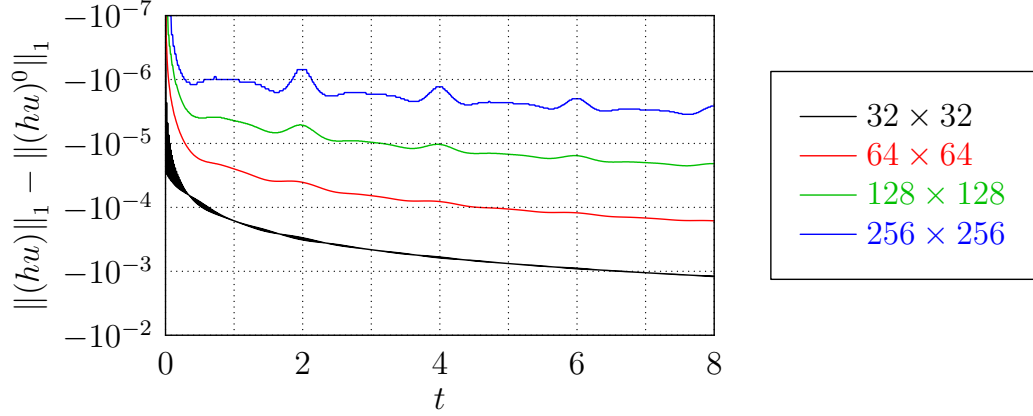
Figure 4.8: Absolute change over time of x-momentum in the discrete L^1 norm defined in (4.29) for $v_{\text{rel}} = -0.5$ and different grid resolutions.

vorticity as stated in (4.24), the vorticity itself changes along fluid columns due to stretching and squeezing induced by the sloping topography (Bühler and Jacobson, 2001), i.e., it must decrease when the fluid becomes shallow and increase when entering deeper regions. By mass conservation, this implies that the cross-sectional area of the vortex is proportional to h^{-1}. In the following, it will be investigated if this is also true for the numerical solution. Furthermore, the experimental order of convergence is estimated by comparing the solutions on different grid levels.

The computational domain is given by $\Omega = [0, 4] \times [0, 1]$ with periodic boundary conditions. To consider different configurations, which only differ by a Galilei transformation also in this case, the quantity $x_{\text{rel}} = x - v_{\text{rel}} \cdot t \pmod 4$ is again defined. The bottom topography has the form

$$b(t, x, y) = \tilde{b}(x - v_{\text{rel}} \cdot t \pmod 4) = \tilde{b}(x_{\text{rel}})$$

with

$$\tilde{b}(x_{\text{rel}}) = \begin{cases} a & \text{if } 0 \leq x_{\text{rel}} \leq 1 \\ a \cdot f(2 - x_{\text{rel}}) & \text{if } 1 < x_{\text{rel}} < 2 \\ 0 & \text{if } 2 \leq x_{\text{rel}} \leq 3 \\ a \cdot f(x_{\text{rel}} - 3) & \text{if } 3 < x_{\text{rel}} < 4 \end{cases}$$
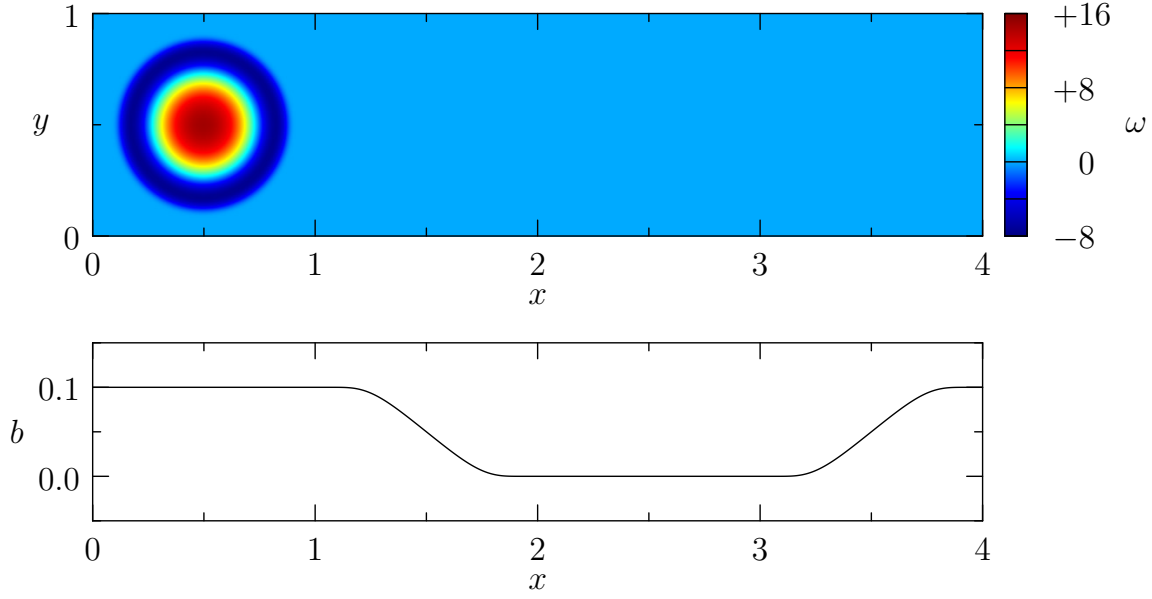
Figure 4.9: Initial vorticity distribution and variation of the bottom topography in x-direction for the "vortex over topography" test case.

and

$$f(x) = \frac{\exp(-1/x)}{\exp(-1/x) + \exp(-1/(1-x))} \ .$$

The function f defines a smooth transition between 0 for $x \searrow 0$ and 1 for $x \nearrow 1$. Note that this function is not defined for $x = 0$ and $x = 1$. The parameter a sets the value of the upper state of the transition and is set to $a = 0.1$ in the presented computations. For $t_{\text{ini}} = 0$, the resulting x-distribution of the bottom topography is displayed in Figure 4.9 (bottom). Two choices of the relative velocity are considered, in particular $v_{\text{rel}} = 0$ and $v_{\text{rel}} = -1$.

The initial momentum is defined by

$$(hu)^0(x,y) = 1 - h^0(x,y)(v_{\text{rel}} + v_\theta(r)\sin\theta) \quad \text{and} \quad (hv)^0(x,y) = h^0(x,y)v_\theta(r)\cos\theta \ ,$$

in which

$$v_\theta(r) = \begin{cases} v_{\text{max}} \dfrac{s \cdot r}{r_m^2 - r^2} \cdot \sqrt{2\exp\left(\dfrac{1}{r^2 - r_m^2}\right)} & \text{for } 0 \leq r < r_m \\ 0 & \text{otherwise} \end{cases} \tag{4.30}$$

91

and

$$r = \sqrt{\left(x - \tfrac{1}{2}\right)^2 + \left(y - \tfrac{1}{2}\right)^2} \quad .$$

In equation (4.30) $v_{\max}$ is the maximum tangential velocity and r_m is the radius of the vortex. In the presented computations, these parameters are set to $v_{\max} = 1.0$ and $r_m = 0.45$. Furthermore, a scaling factor is given by

$$s = \frac{|r_{vm}^2 - r_m^2|}{r_{vm}\sqrt{2\exp(1/(r_{vm}^2 - r_m^2))}} \quad ,$$

where

$$r_{vm} = \frac{1}{2}\sqrt{-2 + 2\sqrt{1 + 4r_m^4}}$$

is the radius at which $v_{\max}$ is attained. The resulting vorticity distribution is displayed in Figure 4.9 (top).

The numerical solution is calculated on different grids with 32×128, 64×256, 128×512 and 256×1024 grid cells. The time step is chosen such that the advective Courant number is roughly cfl $= 0.8$. For $v_{\mathrm{rel}} = 0$ this corresponds to time steps $\Delta t_{32} = 1/100$, $\Delta t_{64} = 1/200$, $\Delta t_{128} = 1/400$ and $\Delta t_{256} = 1/800$. For $v_{\mathrm{rel}} = -1$ the time steps are $\Delta t_{32} = 1/64$, $\Delta t_{64} = 1/128$, $\Delta t_{128} = 1/256$ and $\Delta t_{256} = 1/512$. The solution is computed until $t_{\mathrm{end}} = 3$.

For $v_{\mathrm{rel}} = 0$ details of the computational solution using 256×1024 grid cells are given in Figure 4.10. In this case, the vortex is advected by a background velocity in the x-direction, which ranges between 1 in the deep region and $1/0.9$ in the shallow region. The bottom topography is stationary. The vortex is slightly compressed in the x-direction when it enters deeper waters ($t = 1$). However, at later times this perturbation leads to the breakup of the vortex. In Figure 4.11 it is visualized, how the solution behaves for different grid resolutions. Here, a detail of the solution at time $t = 2$ is displayed. The results clearly suggest convergence to the unknown exact solution.

The corresponding results for $v_{\mathrm{rel}} = -1$ are given in Figures 4.12 and 4.13. In this case, the vortex is roughly stationary, but the bottom topography is advected with velocity -1 in the x-direction. Apart from the Galilei transformation the results for 256×1024 are almost indistinguishable from those with $v_{\mathrm{rel}} = 0$. Only on coarser resolutions differences become visible.
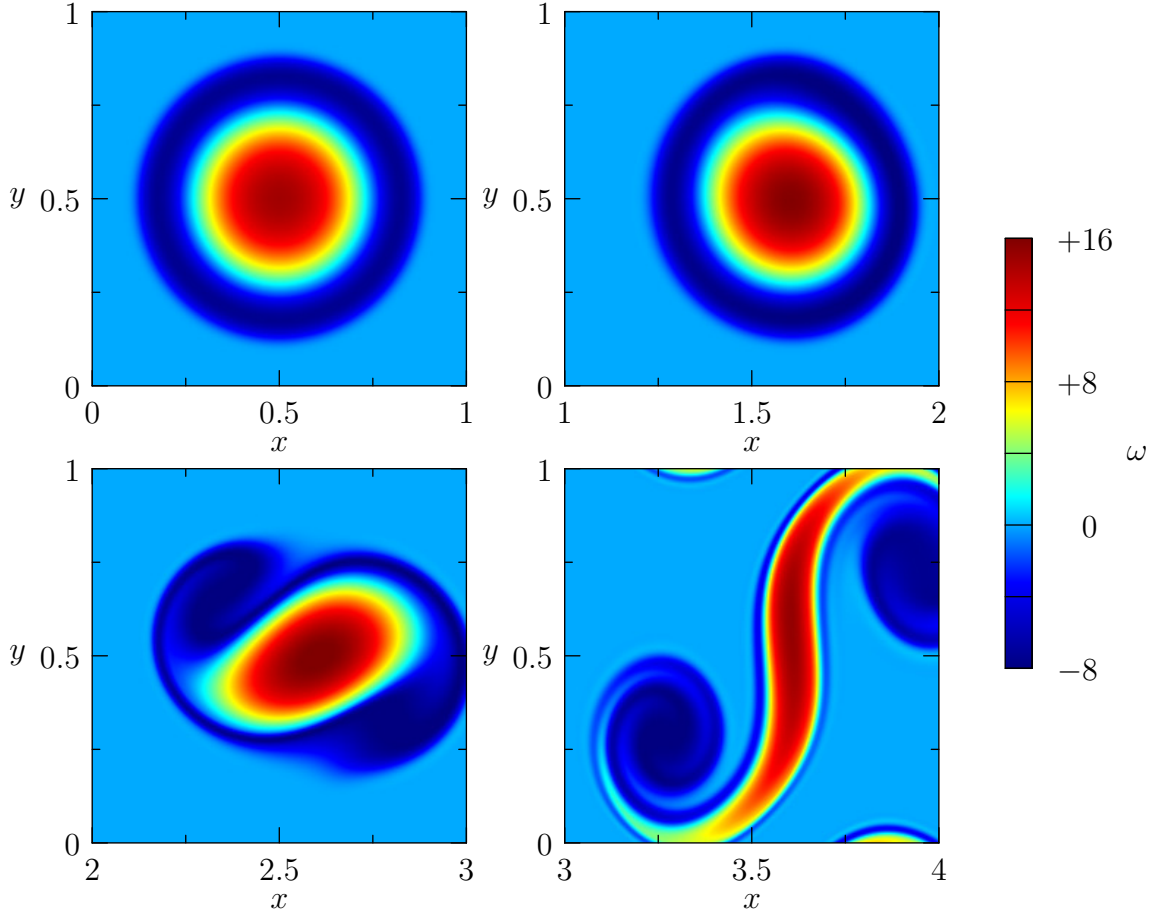
Figure 4.10: Vorticity distribution for the "vortex over topography" test case with $v_{\mathrm{rel}} = 0$ and 256×1024 grid cells. Details of the computational domain at $t_{\mathrm{ini}} = 0$ (top left), $t = 1$ (top right), $t = 2$ (bottom left), $t_{\mathrm{end}} = 3$ (bottom right).

To analyze the relationship between bottom topography and vorticity in the numerical results, the evolution of the global maximum of vorticity is computed. In Figures 4.14 and 4.15, the maximum of the vorticity normalized to the initial fluid depth at the vortex center, i.e. $\max(\omega_{\mathrm{rel}})$ with $\omega_{\mathrm{rel}} = 0.9 \cdot \omega / \max(\omega^0)$, is shown for $v_{\mathrm{rel}} = 0$ and $v_{\mathrm{rel}} = -1$, respectively. Especially for fine grid resolutions, the evolution of this quantity seems to resemble the proportional relationship between ω and $h = 1 - b$ for exact solutions. However, the position of the maximum was not computed in order to further quantify the results. If the resolution becomes too coarse, the numerical errors seem to introduce some vorticity over time.
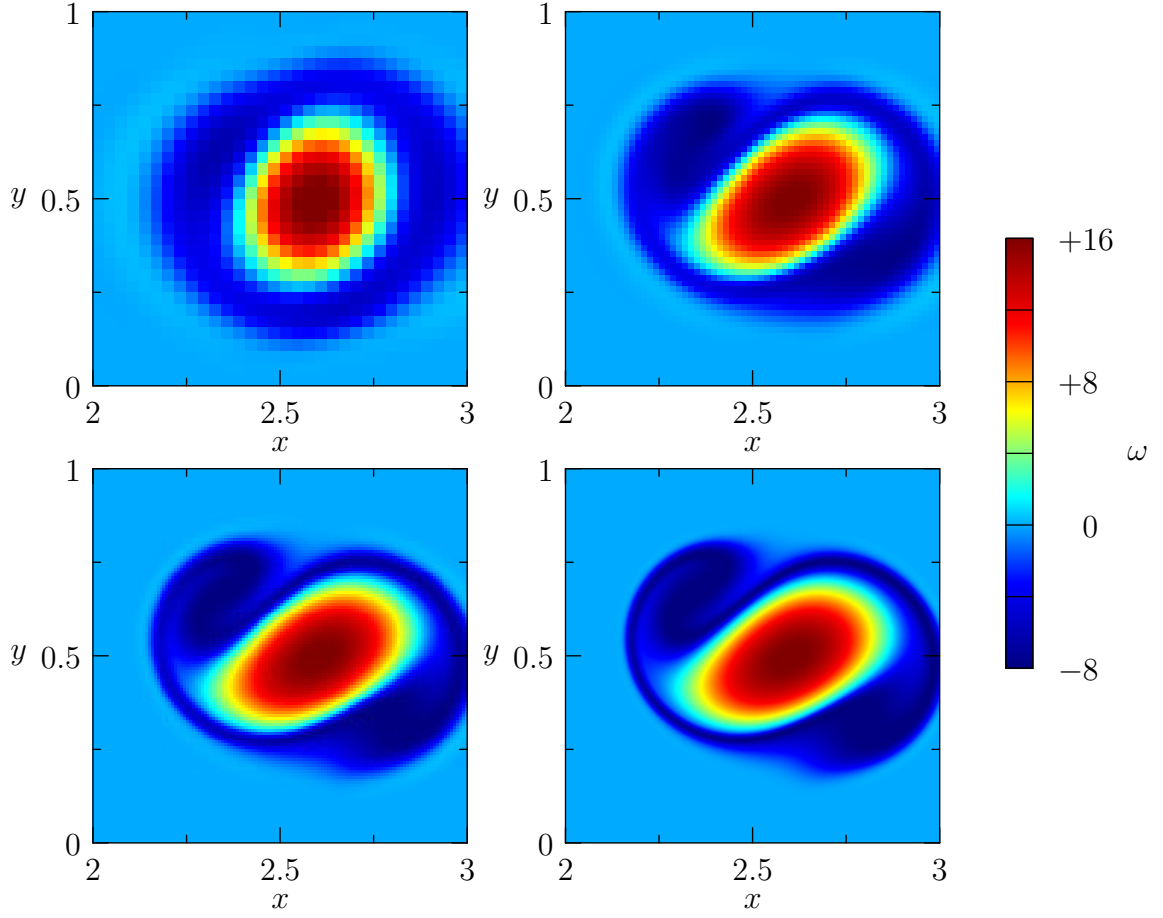
Figure 4.11: Vorticity distribution for the "vortex over topography" test case with $v_{\mathrm{rel}} = 0$ at $t = 2$. Details of the computational domain using 32×128 (top left), 64×256 (top right), 128×512 (bottom left) and 256×1024 (bottom right) grid cells.

Finally, the experimental order of convergence is estimated by comparing the solutions on different grid resolutions at time $t = 2$. For this purpose, a "fine grid" solution is prolongated to the next coarser grid by averaging over the four grid cells which form a cell on the coarse grid, i.e.,

$$
(h\boldsymbol{u})^f_{2i+3/2,2j+3/2} = \frac{1}{4} \left((h\boldsymbol{u})^f_{2i+1,2j+1} + (h\boldsymbol{u})^f_{2i+2,2j+1} + (h\boldsymbol{u})^f_{2i+2,2j+1} + (h\boldsymbol{u})^f_{2i+2,2j+2} \right) .
$$

Then the estimated "error vector" is formed by taking the Euclidean norm of the
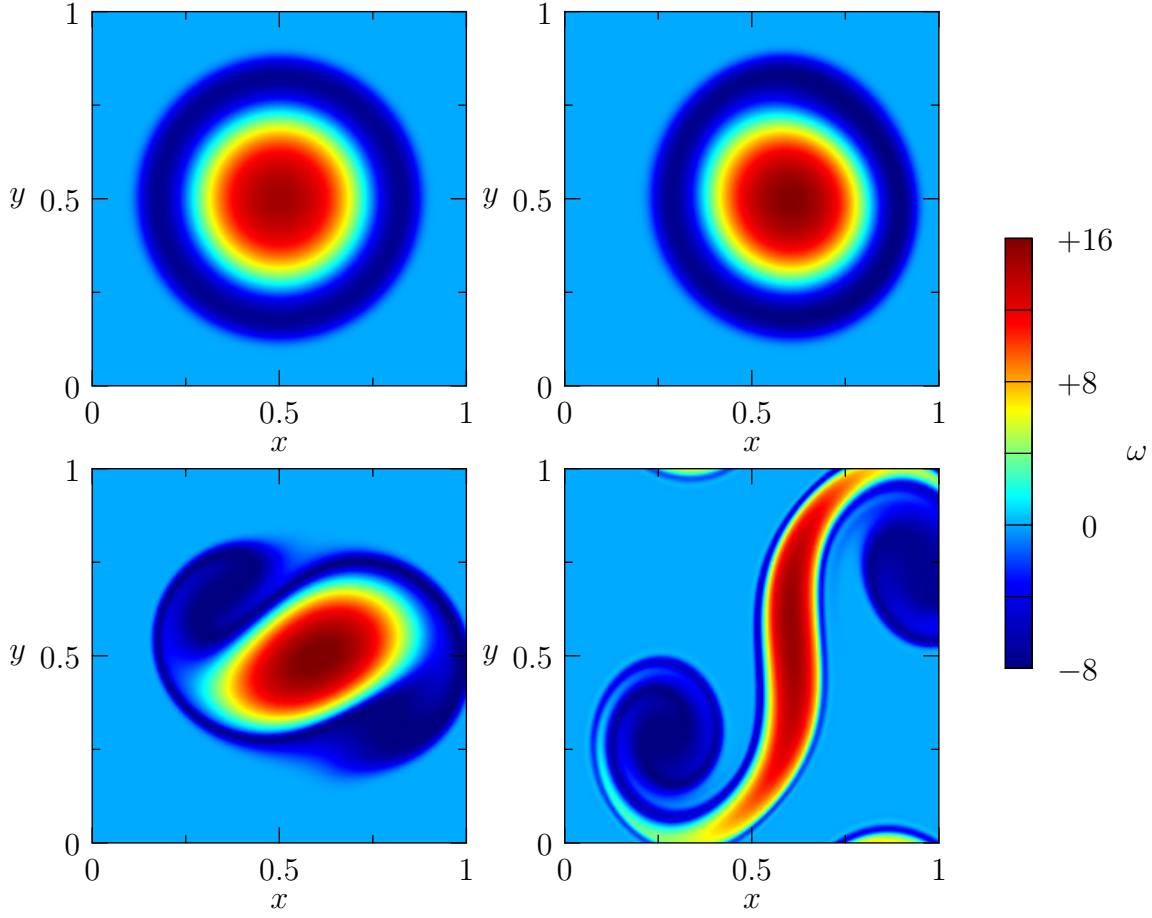
Figure 4.12: Same as Figure 4.10, but with $v_{\mathrm{rel}} = -1$.

difference in momentum from the coarse grid solution and the coarsened fine grid solution:

$$\boldsymbol{e}^c_{2i+3/2,2j+3/2} \coloneqq \left\| (h\boldsymbol{u})^c_{2i+3/2,2j+3/2} - (h\boldsymbol{u})^f_{2i+3/2,2j+3/2} \right\|_2 \ .$$

The convergence rate γ^f_c is then calculated as in (4.28) using the L^2 and the L^∞ norms. The computed rates are summarized in Table 4.1, which include the result obtained from an even finer grid with 512×2048 grid cells (and also a smaller time step). These rates also suggest second-order convergence of the scheme for both choices of v_{rel}, although for such an estimation of the convergence rate the results have still some uncertainty.
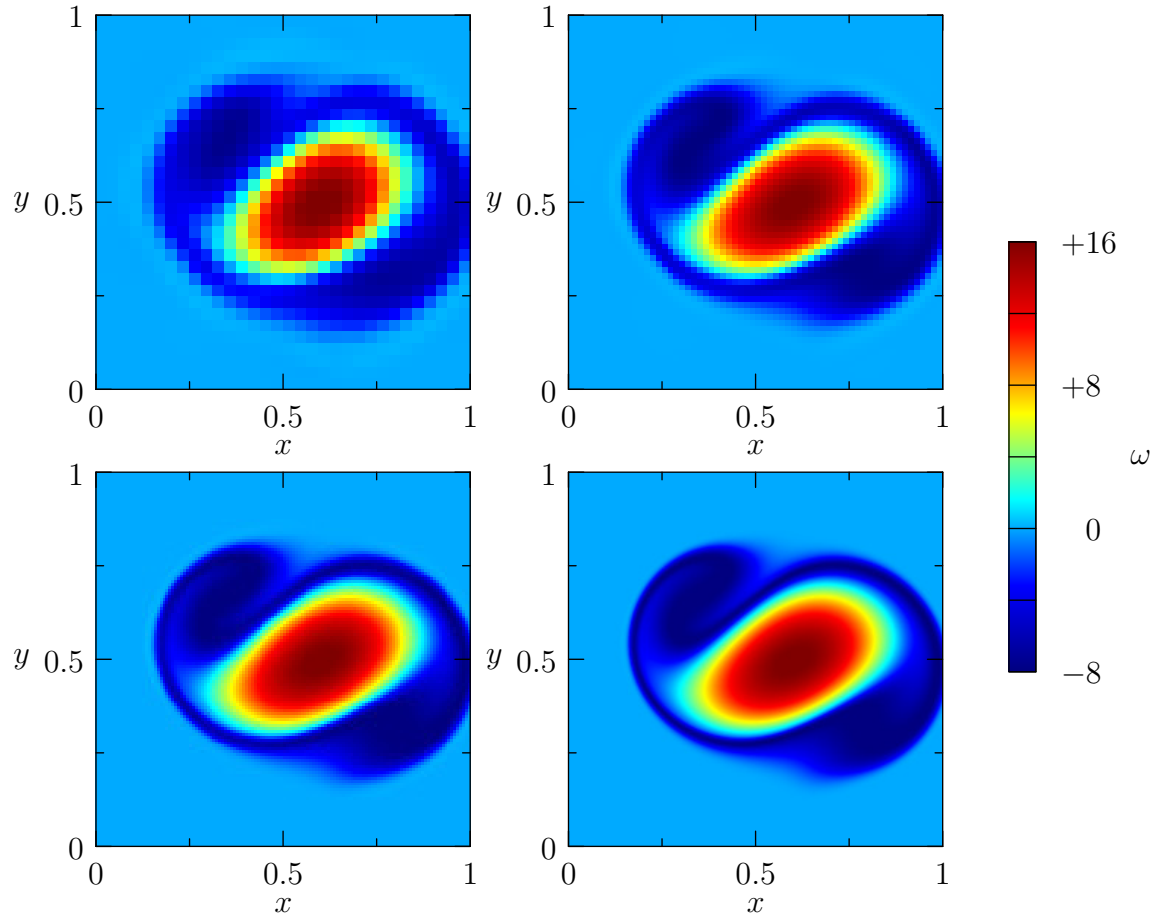
Figure 4.13: Same as Figure 4.11, but with $v_{\mathrm{rel}} = -1$.

Table 4.1: Experimental convergence rates for the "vortex over topography" test case for different choices of v_{rel} at time $t = 2$.

test case	norm	rate γ_{32}^{64}	rate γ_{64}^{128}	rate γ_{128}^{256}
$v_{\mathrm{rel}} = 0.0$	L^2	1.578	2.508	2.406
	L^∞	1.430	2.360	1.993
$v_{\mathrm{rel}} = -1.0$	L^2	2.297	2.365	1.923
	L^∞	1.873	2.015	1.865
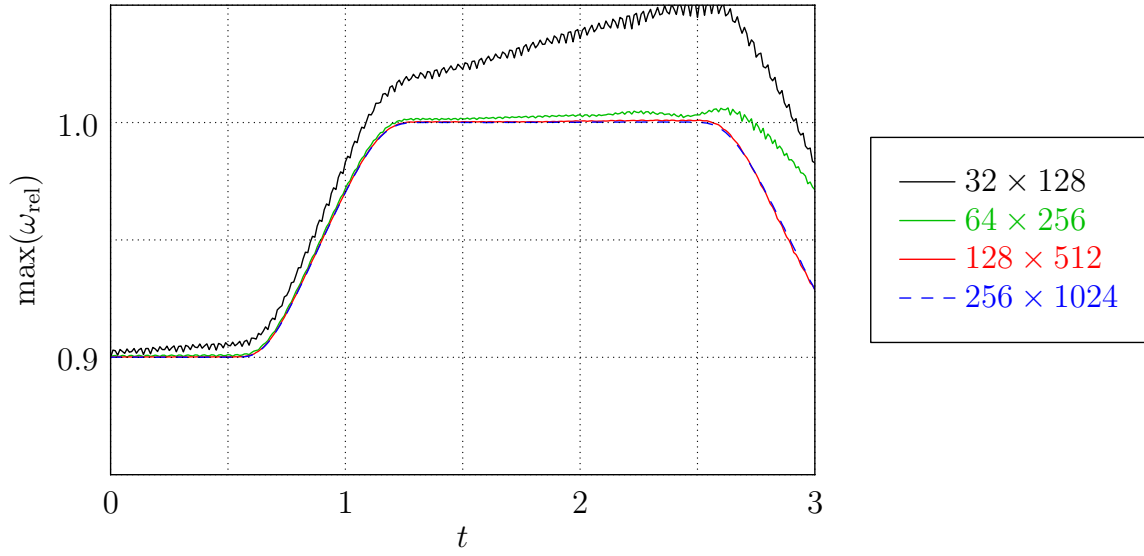
Figure 4.14: Evolution of the maximum in relative vorticity for the "vortex over topography" test case with $v_{\mathrm{rel}} = 0$ for different grid resolutions.
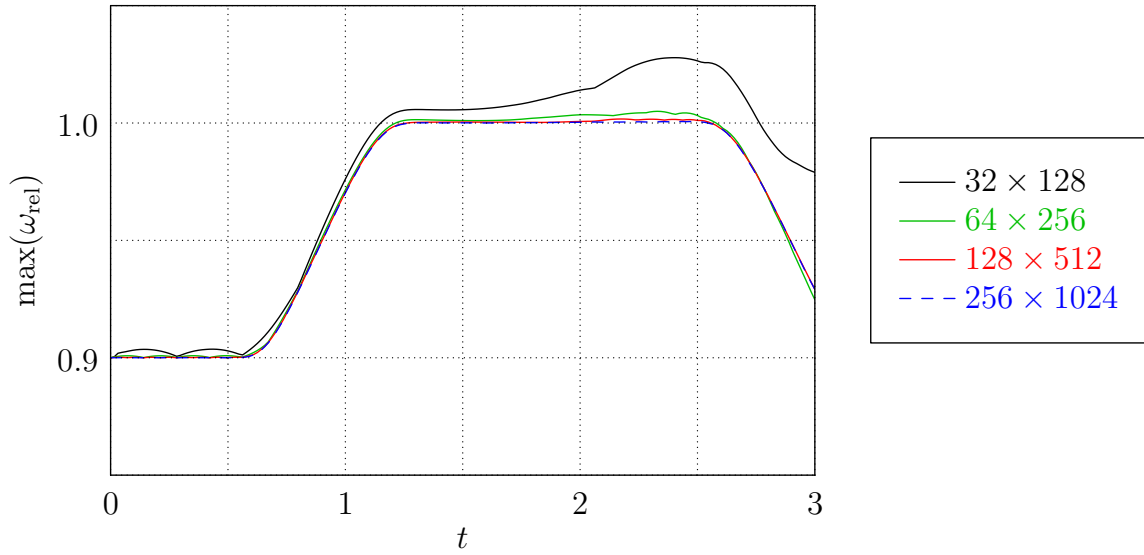
Figure 4.15: Evolution of the maximum in relative vorticity for the "vortex over topography" test case with $v_{\mathrm{rel}} = -1$ for different grid resolutions.

5 Semi-implicit solution of low Froude number shallow water flows

As stated in the introduction, the ultimate goal of this thesis is to develop a semi-implicit numerical scheme for the solution of the fully nonlinear shallow water equations at low Froude numbers, which is of second-order accuracy. This method should be able to correctly reproduce the asymptotic regime of fast gravity waves traveling over rough topography. To achieve this goal, the methods developed in the previous chapters are further extended and combined, leading in the end to a multiscale numerical scheme for the solution of the governing equations.

In a first step, the zero Froude number projection method from Chapter 4 is modified to obtain a semi-implicit method for low Froude number shallow water flows. This method essentially employs a trapezoidal time discretization for the linear implicit part of the method. To be able to combine this method with the multilevel method for the linearized shallow water equations derived in Chapter 3, two variants of the semi-implicit scheme using BDF(2) time discretizations in the linear part are developed. Finally, all ingredients are combined to obtain the multiscale scheme for the shallow water equations at low Froude numbers.

5.1 From zero to low Froude numbers

To extend the projection method from the last chapter to non-zero Froude numbers, the shallow water equations with time dependent bottom topography must be put into a similar form as the Lake equations. As derived in Chapter 2, they are given by

$$
\begin{aligned}
h_t + \nabla \cdot (h\boldsymbol{u}) &= 0 \\
(h\boldsymbol{u})_t + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u}) + \frac{1}{\mathsf{Fr}^2} h \nabla (h + b) &= 0 \ .
\end{aligned}
\tag{5.1}
$$

To reformulate this system, let us assume that the height can be decomposed into

$$h(t, \boldsymbol{x}; \mathsf{Fr}) = h_0(t, \boldsymbol{x}) + \mathsf{Fr}^2\, h'(t, \boldsymbol{x}) \tag{5.2}$$

with

$$h_0(t, \boldsymbol{x}) = H_0(t) - b(t, \boldsymbol{x})\ .$$

Here, H_0 is the mean background total height, which can only change due to flow over the boundary of the domain or to a change in the mean bottom topography. Therefore, h_0 can only change due to boundary flow or (local) change of bottom topography. The dynamics of the flow are thus given by h'.

This ansatz is justified by the asymptotic analysis of the zero Froude number limit of the governing equations. Plugging it into the above system (5.1), the shallow water equations can be rewritten as

$$\begin{aligned} h_t + \nabla \cdot (h\boldsymbol{u}) &= 0 \\ (h\boldsymbol{u})_t + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u}) + h\nabla h' &= 0\ . \end{aligned} \tag{5.3}$$

Compared to the zero Froude number equations, h' takes the role of $h^{(2)}$, but is no longer a Lagrange multiplier. Therefore, also the velocity no longer satisfies a strict divergence constraint. However, at low Froude numbers, these fields should be close to their zero Froude number counterparts. This is due to the mathematical equivalence of the shallow water and the Euler equations and related convergence results for the low Mach number limit of the Euler equations (e.g. Klainerman and Majda, 1981).

5.1.1 Semi-implicit time discretization

Similar to the formulation of the zero Froude number projection method, the semi-implicit scheme is derived by a semi-discretization in time. The discretization in space is discussed in a second step. Following the derivation of the projection method from Chapter 4, the scheme consists of a predictor step, which is followed by two elliptic corrections. The essential difference is that the above ansatz for the height involves the introduction of local time derivatives of this variable. This leads to the solution of (uncritical) Helmholtz problems in the correction steps.

For the sake of completeness, the whole derivation of the scheme is repeated and the differences to the zero Froude number case are pointed out where they occur. The auxiliary system solved in the predictor step, is given by

$$
\begin{aligned}
h_t + \nabla \cdot (h\boldsymbol{u}) &= 0 \\
(h\boldsymbol{u})_t + \nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u}) + \Phi h \nabla (h + b) &= -(1 - \Phi\mathsf{Fr}^2)(h\nabla h')^{\text{old}} \ ,
\end{aligned}
\tag{5.4}
$$

where the right hand side of the momentum equation is given as a "source term" from an old (known) time level. The quantity $\Phi \in \mathbb{R}_0^+$ is again a given parameter of the scheme. Note that for the choice $\Phi = 1/\mathsf{Fr}^2$ the source term would vanish. In this case the auxiliary system is identical to the system we actually want to solve, and there would be no benefit concerning the maximum allowed time step (controlled by the largest characteristic velocity of the system). For this reason it is desirable to have $\Phi \ll 1/\mathsf{Fr}^2$, which slows down the gravity wave speed of the homogeneous system on the left hand side of (5.4).

At the same time, the auxiliary system is chosen in such a way, that by choosing $(h\nabla h')^{\text{old}}(\boldsymbol{x}) = (h\nabla h')(t_0, \boldsymbol{x})$ and starting with the same initial conditions for the governing equations and the auxiliary system, the difference in tendency is

$$
\frac{\partial}{\partial t}[(h\boldsymbol{u}) - (h\boldsymbol{u})^*](x, t_0) = 0 \ ,
$$

where the star denotes the solution of the auxiliary system. Therefore, the solution of the auxiliary system deviates from the one of the shallow water equations by an order of $(t - t_0)^2$ as $t \to t_0$. As in the zero Froude number case, from now on $\Phi = 0$ is chosen if not stated otherwise. This corresponds again to solving the pressureless equations introduced in Section 2.5 in the predictor step. Furthermore, the "source term" of the auxiliary system is set to $(h\nabla h')^{\text{old}}(\boldsymbol{x}) \coloneqq (h\nabla h')(t^n, \boldsymbol{x})$.

Integrating the governing equations from time level t^n to $t^{n+1} \coloneqq t^n + \Delta t$ and evaluating the flux terms at the half-time levels $t^{n+1/2} \coloneqq t^n + \Delta t/2$ yields

$$
h^{n+1} = h^n - \Delta t \left[\nabla \cdot (h\boldsymbol{u})^{n+1/2} \right]
\tag{5.5}
$$

and

$$
(h\boldsymbol{u})^{n+1} = (h\boldsymbol{u})^n - \Delta t \left[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2} + (h\nabla h')^{n+1/2} \right] \ ,
\tag{5.6}
$$

which is second-order accurate. Here and in the following $h(\boldsymbol{x}, t^n)$ is abbreviated by h^n etc. Once again, a height correction $\delta h_{\text{fl}}^{',n}$ is introduced, and the momentum (and similarly the velocity) at the half time level is computed by

$$(h\boldsymbol{u})^{n+1/2} = (h\boldsymbol{u})^{*,n+1/2} - \frac{\Delta t}{2} h^n \nabla \delta h_{\text{fl}}^{',n} \ , \tag{5.7}$$

where $(h\boldsymbol{u})^{*,n+1/2}$ is the momentum computed by the auxiliary system. The divergence of this equation in combination with the height update (5.5) leads to the first correction equation, which is

$$-\frac{\mathsf{Fr}^2}{\Delta t} \delta h_{\text{fl}}^{',n} + \frac{\Delta t}{2} \nabla \cdot (h^n \nabla \delta h_{\text{fl}}^{',n}) = \frac{H_0^{n+1} - H_0^n}{\Delta t} - \frac{b^{n+1} - b^n}{\Delta t} - \frac{h^{*,n+1} - h^{*,n}}{\Delta t} \ . \tag{5.8}$$

This is an (uncritical) Helmholtz problem for $\delta h_{\text{fl}}^{',n}$. The last term on the right hand side is obtained by substituting the auxiliary momentum through the height equation of (5.4). Note that for $\mathsf{Fr} = 0$ this equation becomes identical to the first correction (4.8) of the projection method in Chapter 4. Using (5.7), the height at the new time level as given in (5.5) and the advective components of the momentum flux can be computed. This leads to an intermediate momentum update defined by

$$(h\boldsymbol{u})^{**} \coloneqq (h\boldsymbol{u})^n - \Delta t \left[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2} + (h\nabla h')^n \right] \ , \tag{5.9}$$

and the momentum at the new time level can be calculated by

$$(h\boldsymbol{u})^{n+1} = (h\boldsymbol{u})^{**} - \frac{\Delta t}{2} \left(\delta h^n \nabla h^{',n} + h^{n+1/2} \nabla \delta h^{',n} \right) \ , \tag{5.10}$$

where $\delta h^n \coloneqq h^{n+1} - h^n$ and $h^{n+1/2} \coloneqq \frac{1}{2}(h^n + h^{n+1})$.

As in the zero Froude number case, a trapezoidal time discretization of the height equation is used to derive an equation for the second correction. It is given by

$$\frac{1}{2} \left[\nabla \cdot (h\boldsymbol{u})^{n+1} + \nabla \cdot (h\boldsymbol{u})^n \right] = -\frac{h^{n+1} - h^n}{\Delta t} \ . \tag{5.11}$$

By substitution of (5.10) into this equation, we finally obtain a second (uncritical)

Helmholtz equation for the height update $\delta h'^{,n} = h'^{,n+1} - h'^{,n}$:

$$-\frac{2\mathsf{Fr}^2}{\Delta t}\delta h'^{,n} + \frac{\Delta t}{2}\nabla \cdot (\widehat{h}^{n+1/2}\nabla \delta h'^{,n}) = 2\frac{H_0^{n+1} - H_0^n}{\Delta t} - 2\frac{b^{n+1} - b^n}{\Delta t}$$
$$+ \nabla \cdot (h\boldsymbol{u})^n + \nabla \cdot (h\boldsymbol{u})^{**} - \frac{\Delta t}{2}\nabla \cdot (\widehat{\delta h}^n\nabla h'^{,n}) \,.$$

(5.12)

Apart from the last term on the right hand side, for $\mathsf{Fr} = 0$ this equation is again essentially equivalent to the zero Froude number case (4.12). In case of the zero Froude number projection method, this last term (without the hat over δh^n) appears in the intermediate momentum update (4.9), since there the height update is given through $H_0(t)$ and $b(t, \boldsymbol{x})$. In the low Froude number case, however, we have $\delta h^n = \delta H_0^n + \mathsf{Fr}^2\delta h'^{,n}$, which means that the part $\mathsf{Fr}^2\Delta t/2\nabla \cdot (\delta h'^{,n}\nabla h'^{,n})$ should be on the left hand side of the equation, modifying the solution operator. This issue is solved by using the height update known from the first correction (denoted by the hat) to compute the right hand side of the Helmholtz equation. The same is true for the weight of the Laplacian in the Helmholtz operator of the left hand side. However, for the final momentum update (5.10), the solution $\delta h'^{,n}$ of (5.12) must be used to compute δh^n in order to get conservation of momentum in the absence of bottom topography.

Remark 5.1 *Since the discretized time derivative of the height perturbation (first term in equation (5.12)) only has a small contribution of order Fr^2 in the calculation of $\delta h'^{,n}$, it could be argued that this term can be also approximated by the result of the first correction equation (5.8) and therefore be written on the right hand side of the equation. This leads to a Poisson-type equation for $\delta h'^{,n}$. However, the numerical results derived with this modification did not deviate much from those obtained by solving the Helmholtz problem (5.12) (not shown), and therefore this approach was not further pursued.* ◁

Remark 5.2 *While it is no issue from the mathematical point of view, using floating point arithmetic one cannot update h' separately from h by using*

$$h'^{,n+1} = h'^{,n} + \delta h'^{,n}$$

and (5.5) at the same time, as done in the zero Froude number case for h and $h^{(2)}$. Here, h and h′ are directly linked through (5.2), and using different equations for the update would potentially result in a violation of this relationship after a number of time steps. Therefore, at the beginning of each time step, h′ is computed from h by using this relation, i.e.,

$$h'^{,n} = \frac{1}{\mathsf{Fr}^2} \left(h^n - H_0^n + b^n \right) \ .$$
(5.13)

◁

5.1.2 Space discretization

The space discretization for the semi-implicit method is also essentially the same as in the zero Froude number projection method. The major differences are that for non-zero Froude numbers two Helmholtz problems must be solved instead of Poisson-type problems, and that some care needs to be taken in order to get conservation of momentum in the absence of bottom topography.

For the Helmholtz problems the scalar term must be discretized additionally to the weighted Laplacian, the latter being discretized as in the zero Froude number case. Consistent with the finite volume flux divergence used for the discrete Laplacian, the mean over one grid cell is taken for the scalar value. For the second correction equation this is

$$p_{\bar{V}} := \frac{1}{|\bar{V}|} \int_{\bar{V}} p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \ ,$$

which results (due to the piecewise bilinear distribution of p) in a nine-point stencil also for this part of the operator (see Figure 5.1). Note, however, that one could also use just the node value in the center of $\bar{V}$ for the discretization of this term, since this is second-order accurate compared to the above average of p. Also, as it has been already pointed out that the scalar part is weighted with the Froude number. Therefore, it becomes less significant for small Froude numbers. For the first correction, the calculation is simply shifted by half a grid spacing in both coordinate directions.
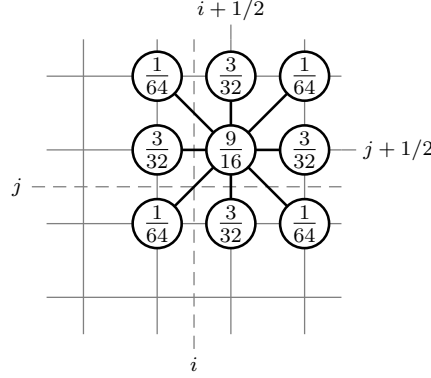
Figure 5.1: Stencil of the scalar part of the Helmholtz operator for a Cartesian grid with uniform grid spacings Δx and Δy in case of the second correction.

Concerning conservation of momentum in case of flat bottom topography, it must be ensured that the term

$$h\nabla h' = h_0 \nabla h' + \mathsf{Fr}^2 h' \nabla h' \tag{5.14}$$

in the momentum equation can be written as a divergence on the discrete level. Since h_0 is constant in this case, this is no problem for the first term on the right hand side of (5.14). For the second term, the equality

$$h' \nabla h' = \frac{1}{2} \nabla (h')^2 \tag{5.15}$$

has to be achieved on the discrete level. At the beginning of each time step h' is derived from cell centered values of the height h. The gradient of the source term in the predictor, however, uses node values. Therefore, h' is first linearly interpolated to the nodes, and then the average of the four surrounding node values in each cell is taken for the factor in front of the gradient. In one space dimension this results in

$$(h'h'_x)_i = \left(\frac{h'_{i+1/2} + h'_{i-1/2}}{2} \right) \left(\frac{h'_{i+1/2} - h'_{i-1/2}}{\Delta x} \right) = \frac{1}{2} \left(\frac{(h'_{i+1/2})^2 - (h'_{i-1/2})^2}{\Delta x} \right) ,$$

which is the discrete equivalent to (5.15).

The whole scheme is implemented for the one- and two-dimensional case. For the one-dimensional version, the spatial operators are just the projections of the two-dimensional case.

5.2 BDF(2)-based semi-implicit discretizations

Implicit time discretizations based on the trapezoidal rule tend to conserve the energy of all wave modes as already described in Chapter 3. However, this is not desirable for high-wave-number modes in this study, due to the unfavorable discrete dispersion relation of such schemes. On the other hand, it was also shown that backward differentiation (BDF) schemes are able to filter these short wave modes in a consistent way. The combination of trapezoidal and BDF time discretizations led to a multilevel scheme for the linearized shallow water equations. To apply this methodology also to the fully nonlinear equations, a semi-implicit scheme has to be derived, which is based on backward differentiation in the linear implicit part of the scheme. In the following, two such discretizations are presented, one where only the second correction equation is discretized using a BDF(2) time discretization, and another one, where both corrections are based on this method. In this work, only uniform time steps are considered. This simplifies the application of multistep methods, since it is not required to account for the different time step sizes. However, often these methods can be generalized to variable time steps as in the case of BDF(2) (Deuflhard and Bornemann, 2002).

5.2.1 BDF(2) in the second correction

By replacing the trapezoidal time discretization (5.11) with a BDF(2) discretization and also doing backward differentiation for the momentum equation, which is

$$(h\boldsymbol{u})^{n+1} = \frac{4}{3}(h\boldsymbol{u})^n - \frac{1}{3}(h\boldsymbol{u})^{n-1} - \frac{2\Delta t}{3}\left[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1} + (h\nabla h')^{n+1}\right],\qquad(5.16)$$

a new second correction equation is obtained. Since this method should be combined with the "standard" scheme derived in Section 5.1 later in the multiscale method, it is desirable to have as few as possible differences between the two variants. Therefore, the discretization of the height equation (5.5) is left untouched, and the discretization of the predictor step and the first correction is the same as in the "standard" scheme. After these two steps, the flux for the intermediate momentum update $(h\boldsymbol{u})^{**}$ (equation (5.9)) is computed, where the advective flux component $\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2}$ is

approximated at the half time level. Since for the BDF discretization this term is needed at the full time level t^{n+1}, it is linearly extrapolated from older time levels, i.e.,

$$(h\boldsymbol{u} \circ \boldsymbol{u})^{n+1} := (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2} + \frac{1}{2}\left((h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2} - (h\boldsymbol{u} \circ \boldsymbol{u})^{n-1/2}\right) \ .$$

The resulting intermediate momentum update is then given by

$$(h\boldsymbol{u})^{**} := \frac{4}{3}(h\boldsymbol{u})^n - \frac{1}{3}(h\boldsymbol{u})^{n-1} - \frac{2\Delta t}{3}\left[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1} + (h\nabla h')^n\right] , \qquad (5.17)$$

and the momentum at the new time level is computed by

$$(h\boldsymbol{u})^{n+1} = (h\boldsymbol{u})^{**} - \frac{2\Delta t}{3}\left(\delta h^n \nabla h'^{,n} + h^{n+1}\nabla \delta h'^{,n}\right) \ . \qquad (5.18)$$

Remark 5.3 *Note that in calculating just an extrapolation for $(h\boldsymbol{u} \circ \boldsymbol{u})$, one first has to subtract the part associated with $(h\nabla h')^n$ from the computed flux prediction of the auxiliary system, then extrapolate, and finally add $(h\nabla h')^n$ again. One could also think of extrapolating the whole sum $(h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2} + (h\nabla h')^n$ to $(h\boldsymbol{u} \circ \boldsymbol{u})^{n+1} + (h\nabla h')^{n+1/2}$. However, this turned out to be a less stable approach.* ◁

To obtain the second correction equation, the momentum update (5.18) is then combined with

$$h^{n+1} = \frac{4}{3}h^n - \frac{1}{3}h^{n-1} - \frac{2\Delta t}{3}\left[\nabla \cdot (h\boldsymbol{u})^{n+1}\right]$$

(instead of (5.11)), which is a BDF(2) discretization of the height equation. This leads to the (uncritical) Helmholtz problem

$$-\frac{3\mathsf{Fr}^2}{2\Delta t}\delta h'^{,n} + \frac{2\Delta t}{3}\nabla \cdot (\widehat{h}^{n+1}\nabla \delta h'^{,n}) = -\frac{\mathsf{Fr}^2}{2\Delta t}\delta h'^{,n-1} + \frac{1}{\Delta t}\left(\frac{3}{2}h_0^{n+1} - 2h_0^n + \frac{1}{2}h_0^{n-1}\right) +$$
$$\nabla \cdot (h\boldsymbol{u})^{**} - \frac{2\Delta t}{3}\nabla \cdot (\widehat{\delta h}^n\nabla h'^{,n})$$
$$(5.19)$$

for $\delta h'^{,n}$, where $h_0^n = H_0^n - b^n$. Here again, the values with the hats are estimates obtained from the height computed in the first correction. To conserve momentum in the absence of bottom topography, also in this case the result of (5.19) must be used in the final momentum update (5.18) for the calculation of δh^n and h^{n+1}. The

space discretization is the same as in the standard scheme. The only difference in the second correction equation is that the Helmholtz operator has different weights.

5.2.2 BDF(2) in both corrections

In order to have BDF(2) discretizations in both corrections, the height equation is also discretized by this method, and the values at the new time level are computed by

$$h^{n+1} = \frac{4}{3}h^n - \frac{1}{3}h^{n-1} - \frac{2\Delta t}{3}\left[\nabla \cdot (h\boldsymbol{u})^{n+1}\right] \tag{5.20}$$

and

$$(h\boldsymbol{u})^{n+1} = \frac{4}{3}(h\boldsymbol{u})^n - \frac{1}{3}(h\boldsymbol{u})^{n-1} - \frac{2\Delta t}{3}\left[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1} + (h\nabla h')^{n+1}\right] . \tag{5.21}$$

In this variant of the semi-implicit scheme the first correction is again calculated by the solution of the auxiliary system together with a height increment. However, this time the approximation is done at the full time level, i.e.,

$$(h\boldsymbol{u})^{n+1} = (h\boldsymbol{u})^{*,n+1} - \Delta t\, h^n \nabla \delta h_{\mathrm{fl}}'^{,n} . \tag{5.22}$$

To obtain $(h\boldsymbol{u})^{*,n+1}$, the same approach from the previous section could be followed by extrapolation of the half time level values $(h\boldsymbol{u})^{*,n-1/2}$ and $(h\boldsymbol{u})^{*,n+1/2}$, which are computed (as fluxes) by the predictor. But this ansatz turns out to be unstable in practical computations. Instead, two successive explicit Euler steps with time step Δt are calculated in the predictor step and the flux computation of the height equation from the second Euler step is utilized as an estimate for $(h\boldsymbol{u})^{*,n+1}$. The determining equation for $\delta h_{\mathrm{fl}}'^{,n}$ is derived by taking the divergence of (5.22) in combination with (5.20), and it is given by the (uncritical) Helmholtz equation

$$-\frac{\mathsf{Fr}^2}{\Delta t}\delta h_{\mathrm{fl}}'^{,n} + \frac{2\Delta t}{3}\nabla \cdot (h^n \nabla \delta h_{\mathrm{fl}}'^{,n}) = -\frac{\mathsf{Fr}^2}{3\Delta t}\delta h'^{,n-1} + \frac{2}{3}\nabla \cdot (h\boldsymbol{u})^{*,n+1} +$$
$$\frac{2}{3\Delta t}\left(\frac{3}{2}h_0^{n+1} - 2h_0^n + \frac{1}{2}h_0^{n-1}\right) .$$

The second flux update is then done in the same way as described in Section 5.2.1 for the method with BDF(2) only in the second correction.

Remark 5.4 *By using a BDF-type discretization as in (5.20) and (5.21), the resulting scheme loses the strict conservation property, that is the new time level value in a cell is computed by the cell value at the old time level together with the flux across the boundary of this cell during this time step. However, by recursion one can rewrite the BDF(2) scheme (for example in case of the height equation) as*

$$h^{n+1} = h^n + \frac{1}{3^{n+1}} \left(h^0 - h^{-1} \right) - \sum_{k=0}^{n} \frac{2\Delta t}{3^{k+1}} \left[\nabla \cdot (h\boldsymbol{u})^{n+1-k} \right] .$$

Therefore, up to an additional contribution of the initial values h^0 and h^{-1} the update essentially consists of a flux divergence composed of all "historical" fluxes. ◁

5.3 Multiscale scheme

With the introduction of BDF(2)-based time discretizations, all ingredients are now at hand to apply the multilevel scheme from Section 3.2 as part of a semi-implicit method to the fully nonlinear shallow water equations. In the following, the "standard" scheme with trapezoidal time discretization in the second correction is combined with the scheme which employs a BDF(2)-based second correction from Section 5.2.1. This is the simplest combination of the introduced schemes, since the scale dependent discretization only involves the second correction. Furthermore, as it was shown in Section 3.2.2, a scale splitting of finite volume type, which would be appropriate for the first correction, involves different discretizations of the Laplacian on each grid level. This makes the splitting much more complicated, if not unfeasible.

With the application of the two schemes, two different intermediate momentum updates are available after the first correction. For the trapezoidal time discretization this is (5.9), i.e.,

$$(h\boldsymbol{u})_{\mathrm{TR}}^{**} := (h\boldsymbol{u})^n - \Delta t \left[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1/2} + (h\nabla h')^n \right] ,$$

whereas for the BDF(2)-based discretization the update is given by (see equation (5.17))

$$(h\boldsymbol{u})_{\mathrm{BDF2}}^{**} := \frac{4}{3}(h\boldsymbol{u})^n - \frac{1}{3}(h\boldsymbol{u})^{n-1} - \frac{2\Delta t}{3} \left[\nabla \cdot (h\boldsymbol{u} \circ \boldsymbol{u})^{n+1} + (h\nabla h')^n \right] .$$

With these updates, the right hand sides of the second correction equations (5.12) and (5.19) are given by

$$f_{\text{TR}}^{\delta h'} = -\frac{\Delta t}{2\text{Fr}^2}\left(2\frac{h_0^{n+1} - h_0^n}{\Delta t} + \nabla \cdot (h\boldsymbol{u})^n + \nabla \cdot (h\boldsymbol{u})_{\text{TR}}^{**} - \frac{\Delta t}{2}\nabla \cdot (\widehat{\delta h}^n \nabla h'^{,n})\right) \quad (5.23)$$

and

$$f_{\text{BDF2}}^{\delta h'} = \frac{1}{3}\delta h'^{,n-1} - \frac{2\Delta t}{3\text{Fr}^2}\left(\frac{3h_0^{n+1} - 4h_0^n + h_0^{n-1}}{2\Delta t} + \nabla \cdot (h\boldsymbol{u})_{\text{BDF2}}^{**} - \frac{2\Delta t}{3}\nabla \cdot (\widehat{\delta h}^n \nabla h'^{,n})\right) . \quad (5.24)$$

Here, both correction equations have been normalized, such that the scalar term in the Helmholtz operator has the common weight 1. The Helmholtz operators are then given by

$$A_{\text{TR}} = \text{id} - \frac{\Delta t^2}{4\text{Fr}^2}\nabla \cdot (\widehat{h}^{n+1/2}\nabla) \quad \text{and} \quad A_{\text{BDF2}} = \text{id} - \frac{4\Delta t^2}{9\text{Fr}^2}\nabla \cdot (\widehat{h}^{n+1}\nabla) . \quad (5.25)$$

Compared to the linear case in Chapter 3, the weights $\widehat{h}^{n+1/2}$ and $\widehat{h}^{n+1}$ in the Laplacian are space dependent. This influences the spatial discretization of the scale dependent operator on the different scales as described below. For the semi-discrete case, the multiscale operator is then given by

$$A := \sum_{\nu=0}^{\nu_M} \left(\mu_\nu A_{\text{TR}} + (1 - \mu_\nu)A_{\text{BDF2}}\right)\left(\Pi_\nu^h - \Pi_{\nu-1}^h\right) ,$$

and the second correction equation for the solution of $\delta h'^{,n+1}$ becomes

$$A\,\delta h'^{,n+1} = \sum_{\nu=0}^{\nu_M}\left(\mu_\nu f_{\text{TR}}^{\delta h',(\nu)} + (1 - \mu_\nu)f_{\text{BDF2}}^{\delta h',(\nu)}\right) . \quad (5.26)$$

With the solution of this Helmholtz problem, the momentum at the new time level is computed according to

$$(h\boldsymbol{u})^{n+1} = \sum_{\nu=0}^{\nu_M}\left(\mu_\nu(h\boldsymbol{u})_{\text{TR}}^{n+1,(\nu)} + (1 - \mu_\nu)(h\boldsymbol{u})_{\text{BDF2}}^{n+1,(\nu)}\right) ,$$

where the scale dependent contributions are computed by blending the updates that

would be obtained by either the trapezoidal or the BDF(2) time discretization. They
are given by projecting

$$(h\boldsymbol{u})_{\text{TR}}^{n+1} = (h\boldsymbol{u})_{\text{TR}}^{**} - \frac{\Delta t}{2} \left(\delta h^n \nabla h'^{,n} + h^{n+1/2} \nabla \delta h'^{,n} \right) \tag{5.27}$$

and

$$(h\boldsymbol{u})_{\text{BDF2}}^{n+1} = (h\boldsymbol{u})_{\text{BDF2}}^{**} - \frac{2\Delta t}{3} \left(\delta h^n \nabla h'^{,n} + h^{n+1} \nabla \delta h'^{,n} \right) \tag{5.28}$$

to each scale. As in the "single-scale" schemes, the height itself, i.e., $h^{n+1/2}$, h^{n+1} and
its update δh^n can be computed from the outcome of the first correction (denoted by
the hats) in the Helmholtz problem (equations (5.23), (5.24) and (5.25)). To obtain
conservation of momentum for flat bottom topography, however, the solution of (5.26)
must be used to compute these values in (5.27) and (5.28).

The blending weights for each grid level are determined similar to the linear case.
For simplicity, the gravity-wave speed $c = \sqrt{h}/\text{Fr}$ is estimated by the square root of
the mean height divided by the global Froude number in the conducted numerical
simulations. This means that the grid-CFL number is given by $\text{cfl}_\nu = c\Delta t/\Delta x_\nu$, where
Δx_ν is the grid spacing on the respective grid level. The blending parameter μ_ν is
then computed according to (3.18).

As mentioned above, the space dependent weights must be taken into account
in the spatial discretization of the Laplacian in the multilevel Helmholtz operator.
As in Chapter 3, only two grid levels in one space dimension are considered, where
the splitting at each node is given by $h_{j+1/2} = h_{j+1/2}^{(f)} + h_{j+1/2}^{(c)}$. Then, the Laplacian
evaluated at nodes common to both grids with cell centered weights w_j splits like

$$\begin{aligned}
L_w(h)_{2j+1/2} \\
&= \frac{1}{\Delta x^2} \left(w_{2j+1}(h_{2j+3/2} - h_{2j+1/2}) - w_{2j}(h_{2j+1/2} - h_{2j-1/2}) \right) \\
&= L_w(h)_{2j+1/2}^{(f)} + \frac{1}{\Delta x^2} \left(w_{2j+1}(h_{2j+3/2}^{(c)} - h_{2j+1/2}^{(c)}) - w_{2j}(h_{2j+1/2}^{(c)} - h_{2j-1/2}^{(c)}) \right) \\
&= L_w(h)_{2j+1/2}^{(f)} + \frac{1}{2\Delta x^2} \left(w_{2j+1}(h_{2j+5/2}^{(c)} - h_{2j+1/2}^{(c)}) - w_{2j}(h_{2j+1/2}^{(c)} - h_{2j-3/2}^{(c)}) \right) \;.
\end{aligned}$$

As before, h has been linearly interpolated between grid points on the coarse grid,
i.e., $h_{2j+3/2}^{(c)} = \frac{1}{2}(h_{2j+1/2}^{(c)} + h_{2j+5/2}^{(c)})$. To obtain essentially the same discretization on

each grid level, the (fine grid) weights w_j are approximated by the coarse grid mean values. This results in

$$L_w(h)_{2j+1/2} \approx L_w(h)^{(f)}_{2j+1/2} + 2L_w(h)^{(c)}_{2j+1/2}$$

with

$$L_w(h)^{(c)}_{2j+1/2} = \frac{1}{(2\Delta x)^2}\Big(\frac{w_{2j+1} + w_{2j+2}}{2}(h^{(c)}_{2j+5/2} - h^{(c)}_{2j+1/2}) -$$
$$\frac{w_{2j-1} + w_{2j}}{2}(h^{(c)}_{2j+1/2} - h^{(c)}_{2j-3/2})\Big) \; .$$

At the nodes only belonging to the fine grid, the Laplacian is split into

$$L_w(h)_{2j+3/2}$$
$$= L_w(h)^{(f)}_{2j+1/2} + \frac{1}{\Delta x^2}\left(w_{2j+2}(h^{(c)}_{2j+5/2} - h^{(c)}_{2j+3/2}) - w_{2j+1}(h^{(c)}_{2j+3/2} - h^{(c)}_{2j+1/2})\right)$$
$$= L_w(h)^{(f)}_{2j+1/2} + \frac{1}{2\Delta x^2}\left(w_{2j+2}(h^{(c)}_{2j+5/2} - h^{(c)}_{2j+1/2}) - w_{2j+1}(h^{(c)}_{2j+5/2} - h^{(c)}_{2j+1/2})\right)$$
$$= L_w(h)^{(f)}_{2j+1/2} + \frac{1}{2\Delta x^2}\left((w_{2j+2} - w_{2j+1})(h^{(c)}_{2j+5/2} - h^{(c)}_{2j+1/2})\right) \; .$$

With the same approximation of the (fine grid) weights w_j, this leads to

$$L_w(h)_{2j+1/2} \approx L_w(h)^{(f)}_{2j+1/2} \; ,$$

which means that essentially the same stencil can be used on each grid level.

This finishes the description of the semi-implicit multiscale method for low Froude number shallow water flows. This method was only implemented in one space dimension. However, the extension to two dimensions should be relatively easy, since the base schemes have been already developed. Only the appropriate splitting of height and (especially) momentum fields has still to be defined.

5.4 Numerical results

Having derived the different schemes for computing low Froude number shallow water flows, in the final section of this chapter the performance of these methods is evaluated for various test cases. Besides the goal of numerically verifying the second-order

accuracy of the methods, some test cases of a more qualitative nature are conducted. These particularly address the asymptotic behavior of the methods in the low Froude number regime as described in Section 2.4.

As in the zero Froude number case, there are not many known exact solutions for the shallow water equations. This is in particular true in the case of nontrivial bottom topography. Some solutions are given by Ball (1965), Thacker (1981), Shapiro (1996), Currò (1989) and Iacono (2005). However, these solutions again pose special requirements on the domain or boundary conditions, like a paraboloidal shape of the basin, which goes beyond the scope of this work.

Only test cases with periodic boundary conditions are considered in either one- or two-dimensional domains. With the exception of the last test case, the computations for the BDF(2) and the multiscale schemes are always started with an initial first step by the "standard scheme". By this, enough old time step values can be provided for the BDF(2)-based schemes. As mentioned above, the blending parameter μ_ν is computed according to (3.18). However, the precise values are always given for reference in each test case.

Since the presented schemes are semi-implicit, two Courant numbers (Courant *et al.*, 1928) are considered in most cases. The Courant number concerning the maximum propagation of information is essentially associated with the propagation of gravity waves in the low Froude number case and denoted by cfl_{grav}. Furthermore, the Courant number concerning advective phenomena (which are mainly computed by the explicit predictor) is given by $\text{cfl}_{\text{adv}} \coloneqq \max_{\boldsymbol{x}}(|\boldsymbol{u}(\boldsymbol{x})|)\Delta t/\Delta x$.

5.4.1 Weakly nonlinear gravity wave

The first test case is set up with initial data, which consists of a smooth right running shallow water simple wave in one space dimension with flat bottom topography. This is similar to one of the test cases for the linearized shallow water equations (see Section 3.3.1). However, due to the nonlinearity of the governing equations, a shock develops after some time. While this is probably one of the most simple setups one can think of, it already reveals some interesting properties of the considered numerical schemes: by the use of the method of characteristics, the exact solution is known until the development of a shock, which is useful for a convergence study. The behavior of

the different schemes towards the compressible regime can be also tested, when the exact solution eventually develops a shock. Furthermore, the evolution of long-wave gravity waves can be analyzed, which is relevant for the asymptotic regime described in Section 2.4 and similar to what was investigated in Section 3.3.1 for the linearized equations.

To derive the initial conditions, let us consider the characteristic variables of the shallow water equations. These are given by (see, e.g., Erbes, 1993)

$$p_1 = u - 2c \quad \text{and} \quad p_2 = u + 2c \ ,$$

where $c = \sqrt{h}/\mathsf{Fr}$ is the gravity wave speed. The definition of a background state $h_0 = 1$ leads to $c_0 = 1/\mathsf{Fr}$. Then, the initial gravity wave speed is given by

$$c = c_0 + c' = \frac{1}{\mathsf{Fr}} + c' \ .$$

To obtain a right running simple wave, the left running characteristic is set to $p_1 = \text{const}$. This constant is chosen to obtain a zero background flow, i.e., $p_1 = -2c_0$, which gives the initial velocity field

$$u = 2(c - c_0) = 2c' \ .$$

Therefore, initially the local Froude number ranges from 0 to

$$\mathsf{Fr}_{\max} \approx \frac{u_{\max}}{c_0} = 2\mathsf{Fr} \max_{x \in \Omega}(c'(x)) \ .$$

For the performed simulations the perturbation of the gravity wave speed is set to $c' = \frac{1}{2}\sin(2\pi x)$. The computational domain is defined by the interval $\Omega = [0, 1]$ with 256 grid cells and periodic boundary conditions.

In a first setup, the Froude number is set to $\mathsf{Fr} = 0.1$ and the time step is chosen to be $\Delta t = 0.003$, which is equivalent to initial Courant numbers $\mathrm{cfl}_{\mathrm{adv}} \approx 0.77$ concerning advection and $\mathrm{cfl}_{\mathrm{grav}} \approx 8.83$ concerning the propagation of gravity waves. In Figure 5.2 the initial conditions for height and momentum are given. The solutions of the numerical schemes are given after 40 time steps ($t = 0.12$) in Figure 5.3 (top) and after 100 time steps ($t = 0.3$) in Figure 5.3 (bottom). At these times the wave has
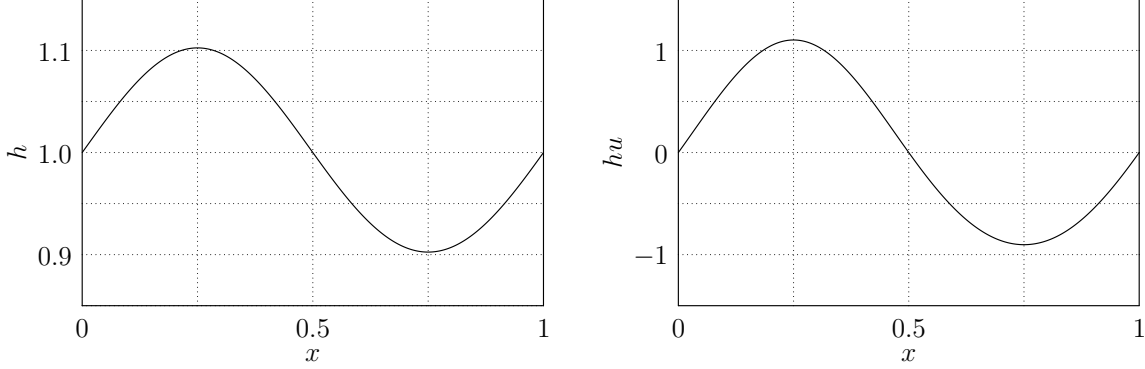
Figure 5.2: Initial conditions for the weakly nonlinear gravity wave test case with Fr $= 0.1$ on a grid with 256 grid cells. Left: height, right: momentum.

traveled approximately 1.2 and 3 times, respectively, through the domain. Since a shock forms at time $t_{\text{shock}} = 1/(3\pi)$, this test shows the performance of the schemes towards the compressible regime. The multiscale scheme is set up with three grid levels and blending factors $\mu_{\boldsymbol{\nu}} = (1, 1/2, 0)$. In addition to the schemes described above, an off-centered version of the standard scheme is implemented. This means that the non-convective flux term $(h\nabla h')^{n+\theta}$ in (5.6) is approximated at $t^n + \theta\Delta t$, and equation (5.11) is substituted by

$$\theta\nabla \cdot (h\boldsymbol{u})^{n+1} + (1-\theta)\nabla \cdot (h\boldsymbol{u})^n = -\frac{h^{n+1} - h^n}{\Delta t} \ , \quad \theta \in [0, 1] \ ,$$

which is also known as θ-scheme. For $\theta = 1$, this method becomes the implicit Euler method. While it is of second-order accuracy only for $\theta = 0.5$ (equivalent to standard scheme), the scheme usually stabilizes for $\theta \in (0.5, 1]$, since more numerical diffusion is introduced. The presented results of this scheme are obtained by setting $\theta = 0.7$.

As one can see in Figure 5.3, for $t = 0.12$ the standard and multiscale schemes develop artificial oscillations in the vicinity of the shock, which continuously grow until either the time step has to be reduced or the schemes become unstable (which already happens before the time $t = 0.3$). Since the initial data only consists of long-wave information, and the contributions on the smaller scales are only small corrections, the results for both schemes are almost identical. On the other hand, the off-centered version of the standard scheme does not show this behavior, and the discontinuity is smeared out by numerical diffusion. The BDF(2)-based schemes show
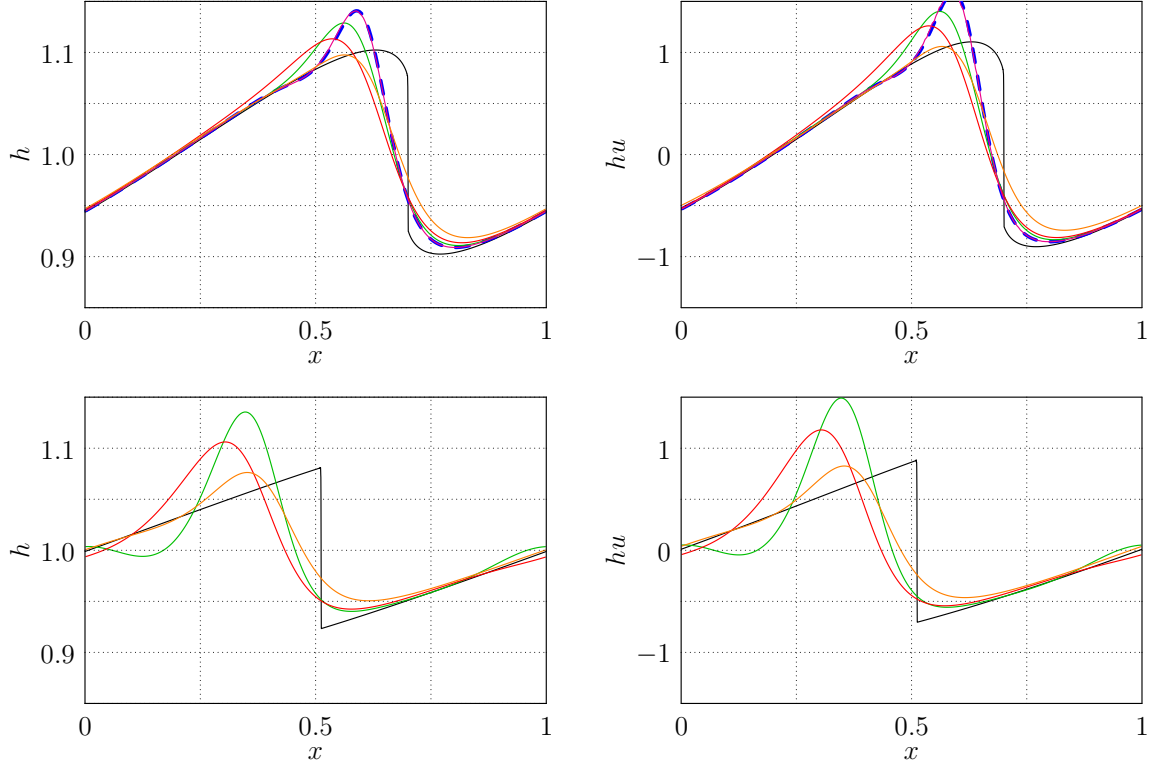
Figure 5.3: Solution of the weakly nonlinear gravity wave test case with $\mathsf{Fr} = 0.1$ at times $t = 0.12$ (top) and $t = 0.3$ (bottom) on a grid with 256 grid cells. Black: exact solution, blue dashed: standard scheme (only for $t = 0.12$), green: BDF(2) in second correction, red: BDF(2) in both corrections, orange: standard scheme off-centered ($\theta = 0.7$), magenta: multiscale scheme.

a behavior which is in between these two extrema, with the version with BDF(2) only in the second correction resulting in a larger overshoot. This becomes even more obvious at the later time $t = 0.3$. Additionally, all schemes introduce a dispersive error in that they slow down the speed of the simple wave.

To test the evolution of long-wave gravity waves the Froude number is reduced to $\mathsf{Fr} = 0.02$ in a second setup. This further decreases the nonlinearity of the equations compared to the case with $\mathsf{Fr} = 0.1$. However, due to the configuration of the initial data, the shock develops at the same time $t_{\text{shock}} = 1/(3\pi)$ as before. The initial conditions for this test case are shown in Figure 5.4 (top). The time step is again $\Delta t = 0.003$, which is equivalent to initial Courant numbers $\text{cfl}_{\text{adv}} \approx 0.77$ and $\text{cfl}_{\text{grav}} \approx 39.55$. The solution at time $t = 0.024$ is displayed in Figure 5.4 (bottom). At
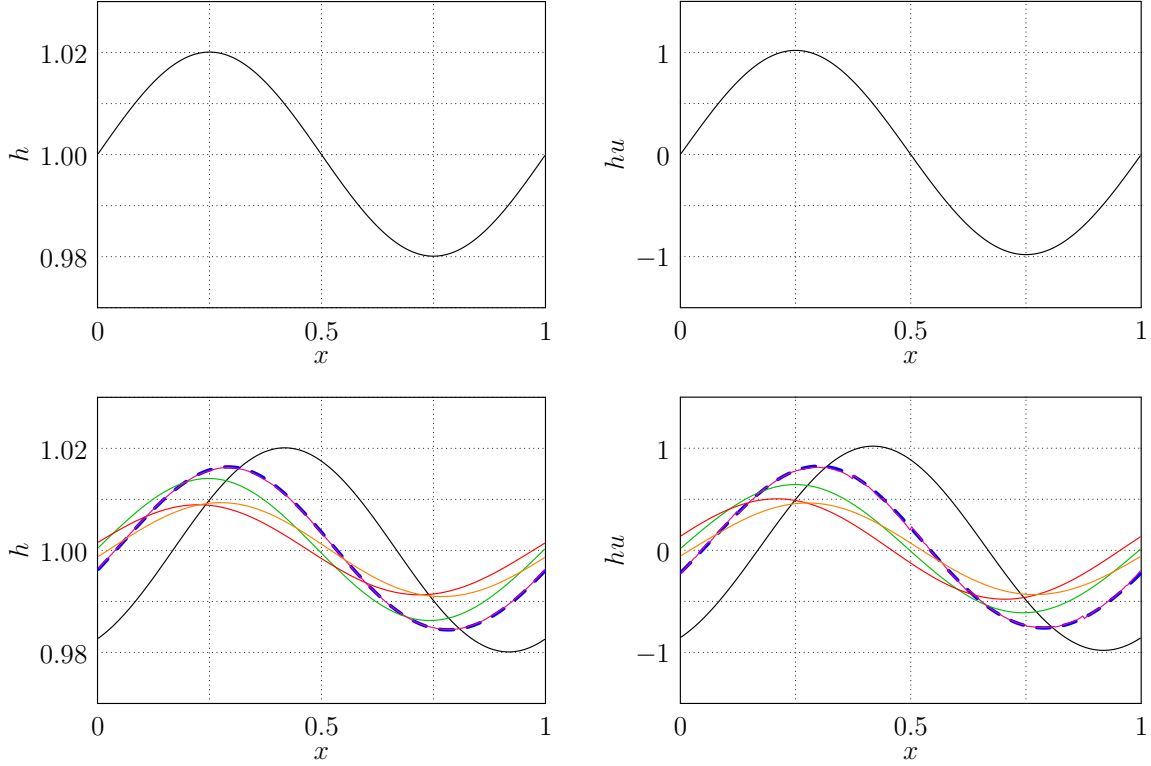
Figure 5.4: Weakly nonlinear gravity wave test case with Fr $= 0.02$ on a grid with 256 grid cells. Initial conditions (top). Solution at $t = 0.024$ (bottom). Black: exact solution, blue dashed: standard scheme, green: BDF(2) in second correction, red: BDF(2) in both corrections, orange: standard scheme off-centered ($\theta = 0.7$), magenta: multiscale scheme.

this time, the gravity wave has traveled approximately 1.2 times through the domain, and its shape has not yet been much distorted compared to the initial data. For this test, the multiscale scheme is applied with six levels and blending parameters $\mu_\nu = (1, 4/5, 3/5, 2/5, 1/5, 0)$.

At the final time the standard and the multiscale scheme show the smallest error in amplitude and phase compared to the exact solution. Also in this case the solutions of these schemes are nearly identical. The worst results are produced by the scheme with BDF(2) discretizations in both corrections. While the amplitude error is comparable to the results derived from the off-centered version of the standard scheme, the phase error is even higher compared to the latter. The method with BDF(2) only in the second correction produces results which are in between these two extrema.

5.4.2 Convergence in one space dimension

The same initial conditions of a right running gravity simple wave and for $\mathsf{Fr} = 0.1$ are used in order to undertake a convergence analysis. The solution of the numerical schemes is computed on different grids and compared to the exact solution at time $t_{\mathrm{end}} = 0.05$. At this time no shock has developed yet, and the true solution can be computed using the method of characteristics. The numerical solution is computed on grids with 256, 512, 1024 and 2048 cells, and the respective time steps are given by $\Delta t_{256} = 1/320$, $\Delta t_{512} = 1/640$, $\Delta t_{1024} = 1/1280$ and $\Delta t_{2048} = 1/2560$. This corresponds to an advective Courant number $\mathrm{cfl}_{\mathrm{adv}} = 0.8$. For the multiscale method, five grid levels with $\mu_{\boldsymbol{\nu}} = (1, 1, 2/3, 1/3, 0)$ are used.

The error of the numerical solutions in the L^{∞} norm is summarized in Figure 5.5. Furthermore, the precise values in the L^2 and L^{∞} norms are given given in the Appendix in Tables A.2 and A.3, where also the convergence rates γ between the grid levels are calculated. The calculation is done in the same way as described in Section 4.2 for the zero Froude number shallow water equations. On fixed grids, the standard scheme produces the smallest errors. The method with a BDF(2)-based second correction produces errors, which are about 1.5 times larger. The biggest errors are produced with the method with BDF(2)-based discretizations in both corrections, which are at most 2.5 larger than those from the standard scheme. The multiscale scheme produces errors, which are comparable with those from the standard scheme. This is again due to the long-wave nature of the initial conditions. As given by the values of μ_{ν}, only the finest scales the BDF(2)-based method is applied, which means that the calculations are nearly identical up to small deviations. The experimental convergence rates suggest for all schemes second-order accuracy.

5.4.3 Convergence in two space dimensions

While in the previous section the convergence test is dominated by gravity wave dynamics, a second convergence test case dominated by advection is discussed in the following. In this test a quasi-stationary vortex is advected in two space dimensions with flat bottom topography.

To derive the initial conditions the shallow water equations are transformed into polar coordinates. For axially symmetric and stationary solutions, this results in the
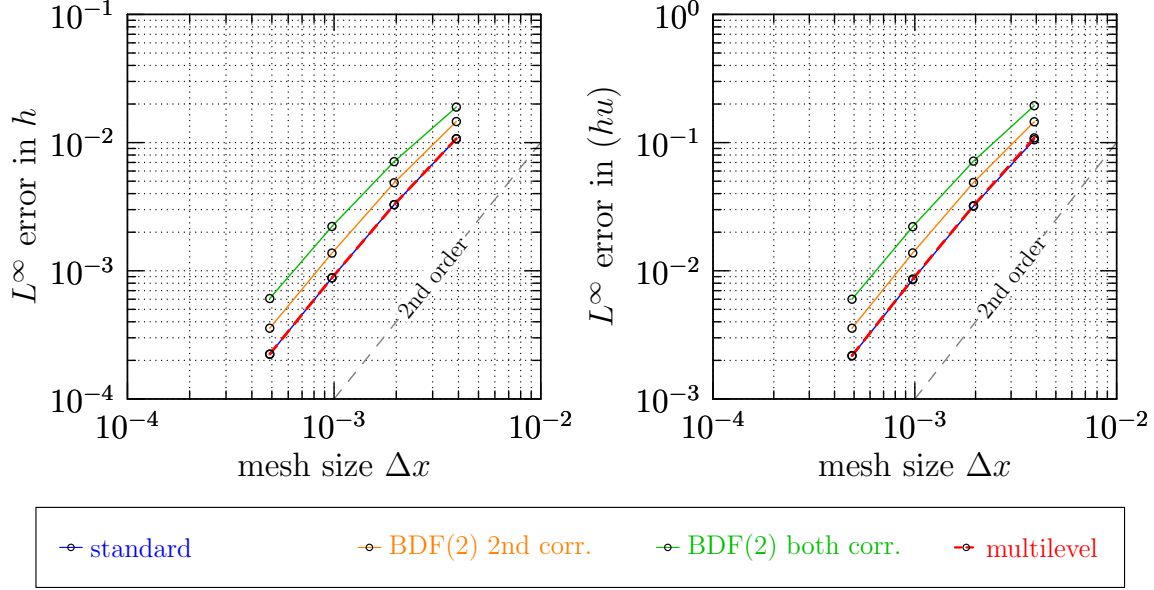
Figure 5.5: L^∞ errors in h and (hu) for the different variants of the semi-implicit method. One-dimensional simple wave test case.

relationship

$$\frac{dh}{dr} = \mathsf{Fr}^2 \frac{v_\theta}{r} \;, \tag{5.29}$$

where r is the radius and $v_\theta(r)$ is the tangential velocity. For the latter, the same structure as in the zero Froude number case is chosen (see Section 4.2.3). It is given by

$$v_\theta(r) = \begin{cases} v_{\max} \dfrac{s \cdot r}{r_m^2 - r^2} \cdot \sqrt{2 \exp\left(\dfrac{1}{r^2 - r_m^2}\right)} & \text{for } 0 \leq r < r_m \\ 0 & \text{otherwise} \;, \end{cases} \tag{5.30}$$

and the radius is computed as distance from the point $(\frac{1}{2}, \frac{1}{2})$:

$$r = \sqrt{\left(x - \tfrac{1}{2}\right)^2 + \left(y - \tfrac{1}{2}\right)^2} \;.$$

The quantity $v_{\max}$ is the maximum tangential velocity within the vortex, and r_m is the radius of the vortex. For the following computations, these parameters are set to
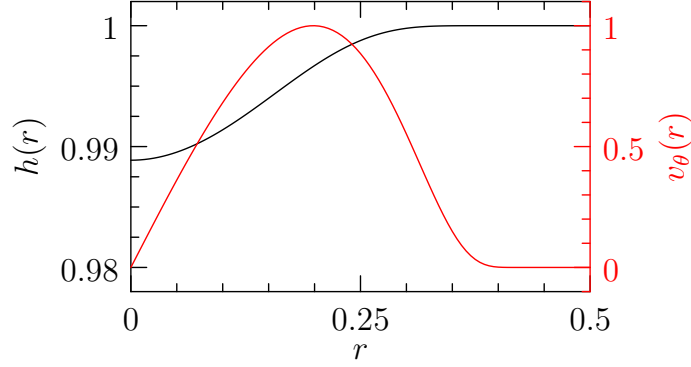
Figure 5.6: Height and tangential velocity as a function of the radius r for a stationary vortex with $v_{\max} = 1.0$, $r_m = 0.45$ and $\mathsf{Fr} = 0.1$.

$v_{\max} = 1.0$ and $r_m = 0.45$. Furthermore, the scaling factor is

$$s = \frac{|r_{vm}^2 - r_m^2|}{r_{vm}\sqrt{2\exp(1/(r_{vm}^2 - r_m^2))}} \ ,$$

where

$$r_{vm} = \frac{1}{2}\sqrt{-2 + 2\sqrt{1 + 4r_m^4}}$$

is the radius at which $v_{\max}$ is attained. Since by (5.29) h is only defined up to a constant, $h_{\mathrm{bg}} = 1$ is chosen as background height. This results in the initial height field

$$h_0(x, y) = \begin{cases} h_{\mathrm{bg}} - v_{\max}s^2\mathsf{Fr}^2\exp\left(\dfrac{1}{r^2 - r_m^2}\right) & \text{for } 0 \le r < r_m \\ h_{\mathrm{bg}} & \text{otherwise} \ . \end{cases} \tag{5.31}$$

The relationship between height and velocity within the vortex is displayed in Figure 5.6. The vortex is advected in the diagonal direction with $(u_{\mathrm{bg}}, v_{\mathrm{bg}}) = (1, 1)$, which defines the initial velocity field

$$u_0(x, y) = u_{\mathrm{bg}} - v_\theta(r)\sin\theta \quad \text{and} \quad v_0(x, y) = v_{\mathrm{bg}} + v_\theta(r)\cos\theta \ , \tag{5.32}$$

with $\theta = \arctan\left((y - \frac{1}{2})/(x - \frac{1}{2})\right)$.

The computational domain is given by the unit square $\Omega = [0, 1]^2$, again with periodic boundary conditions. The numerical solution is calculated on grids with
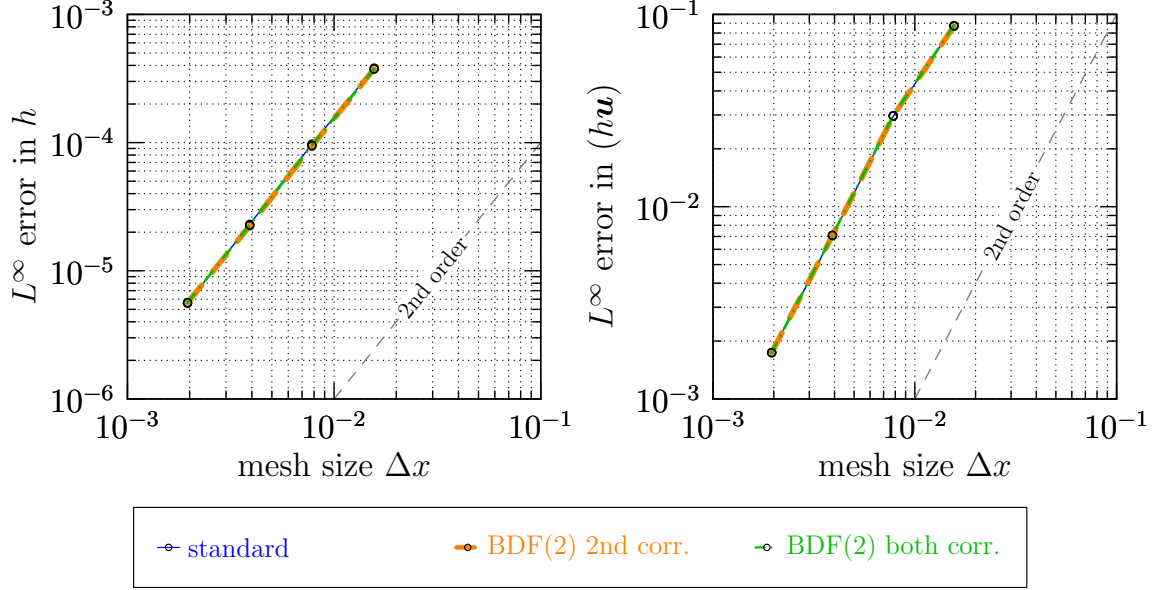
Figure 5.7: L^∞ errors in h and $(h\boldsymbol{u})$ for the different variants of the semi-implicit method. 2D convergence test case.

64×64, 128×128, 256×256 and 512×512 cells. For two space dimensions, the scheme with $\mathrm{BDF}(2)$ in both corrections shows a more restrictive CFL condition in practical computations. Therefore, the time steps for all schemes are set to $\Delta t_{64} = 1/500$, $\Delta t_{128} = 1/1000$, $\Delta t_{256} = 1/2000$ and $\Delta t_{512} = 1/4000$, which is equivalent to an advective Courant number $\mathrm{cfl}_{\mathrm{adv}} \approx 0.44$. The numerical solutions are integrated until the final time $t_{\mathrm{end}} = 1$. At this time, the vortex has traveled once through the domain and the solution is compared to the exact solution, which is identical to the initial data.

The resulting errors and convergence rates are displayed in Figure 5.7 for height and momentum, respectively (see Tables A.4 and A.5 for the detailed results). These are computed in the same way as in Section 4.2. For this test case, the standard and the $\mathrm{BDF}(2)$-based schemes produce almost the same errors in magnitude. This is probably due to the fact that the discretization of the advective terms is essentially the same in all three methods. The experimental convergence rates suggest second-order accuracy also in this case.

5.4.4 Lake at rest

An important property when numerically simulating geophysical flows is the preservation of non-trivial steady states. The simplest one is the so-called "lake at rest" solution, where $\boldsymbol{u} \equiv \boldsymbol{0}$ and $h = C - b$. Since the early works of Gosse and LeRoux (1996) and Greenberg and LeRoux (1996), numerical methods that preserve the "lake at rest" steady state, are usually called well-balanced schemes.

The computational domain consists again of the unit square $\Omega = [0, 1]^2$. As bottom topography the same isolated "hill" is considered, as it was used in case of the zero Froude number equations (see Section 4.2.1). It is defined by

$$b(t, x, y) = \tilde{b}(r) \qquad \text{with} \qquad r = \sqrt{(x - 0.5)^2 + (y - 0.5)^2}$$

and

$$b(r) = \begin{cases} a \cdot \frac{\exp(-0.5/(r_m^2 - r^2))}{\exp(-0.5/r_m^2)} & \text{if } r < r_m \\ 0 & \text{otherwise .} \end{cases}$$

The radius is set to $r_m = 0.3$ and the height of the hill to $a = 0.2$. Initial height and momentum are given by $h(0, x, y) = 1 - b(0, x, y)$ and $(h\boldsymbol{u})(0, x, y) \equiv (0, 0)$, respectively. The computational grid has 128×128 cells, and the time step is chosen to be $\Delta t = 1/100$. The solution is computed until the final time $t_{\text{end}} = 20$.

In Figure 5.8 the evolution of the error in height and momentum is plotted. The error is measured in the L^2 norm as well the supremum (L^∞) norm. After an initial growth during the first 200 to 500 time steps, the error more or less stabilizes at some level, which is prescribed by the accuracy of the linear solver for the correction equations. Only the scheme with BDF(2) in both corrections displays a weak increase over time in the height variable.

Remark 5.5 *Note that the height at the new time level t^{n+1} could be computed in two different ways: Either by computing the flux divergence in (5.5) through (5.7) (in case of the standard scheme), or simply by using the solution of the first correction equation itself, i.e.,*

$$h^{n+1} = h^n + \delta h^n = h^n + (H^{n+1} - H^n) - (b^{n+1} - b^n) + \mathsf{Fr}^2 \delta h'^{,n} \ .$$
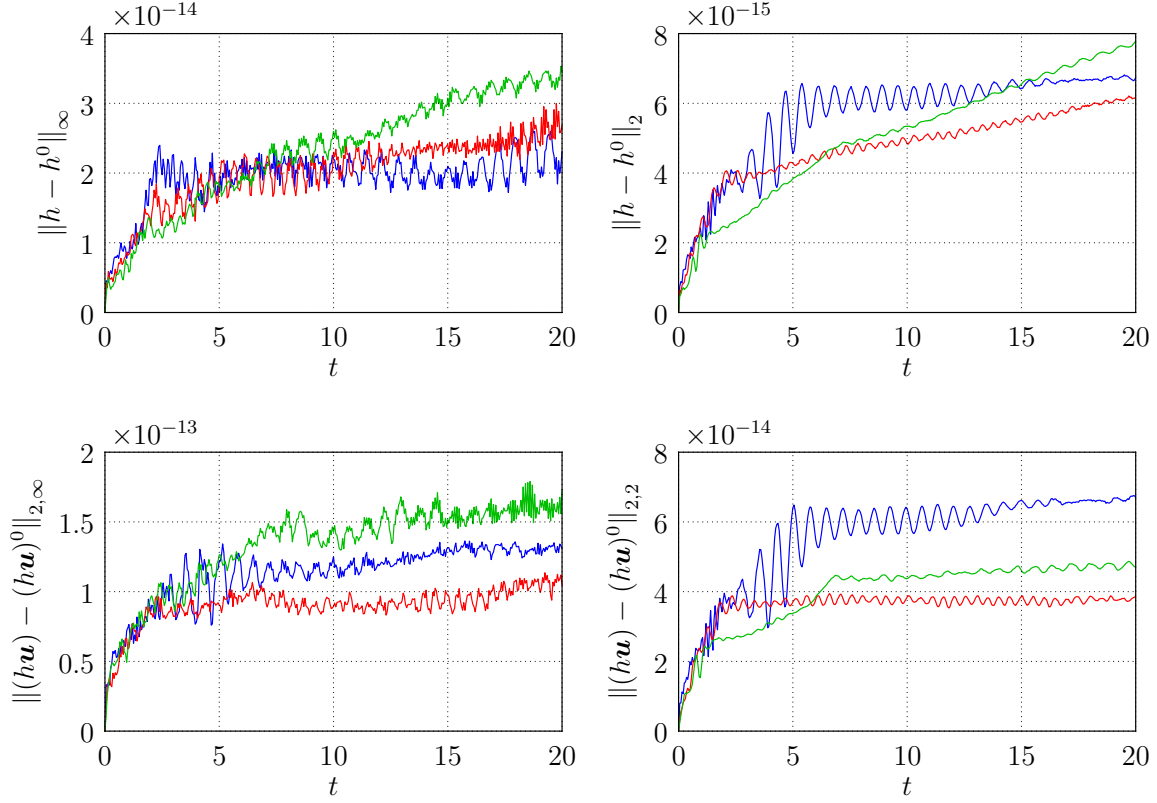
Figure 5.8: Error in height (top row) and momentum (bottom row) for the "lake at rest" test case, measured in the L^2 and L^∞ norms. Blue: standard method, red: BDF(2) in first correction, green: BDF(2) in both corrections.

While the two approaches are mathematically equivalent, in the numerical scheme they lead to slightly different results. In case of the "lake at rest" test case, the latter actually produces better results as can be seen in Figure 5.9. Here, the resulting errors are on the level of machine accuracy. On the other hand, the latter approach produces unstable solutions for other configurations like the advected vortex from the previous test case. Therefore, this strategy was not further investigated. ◁
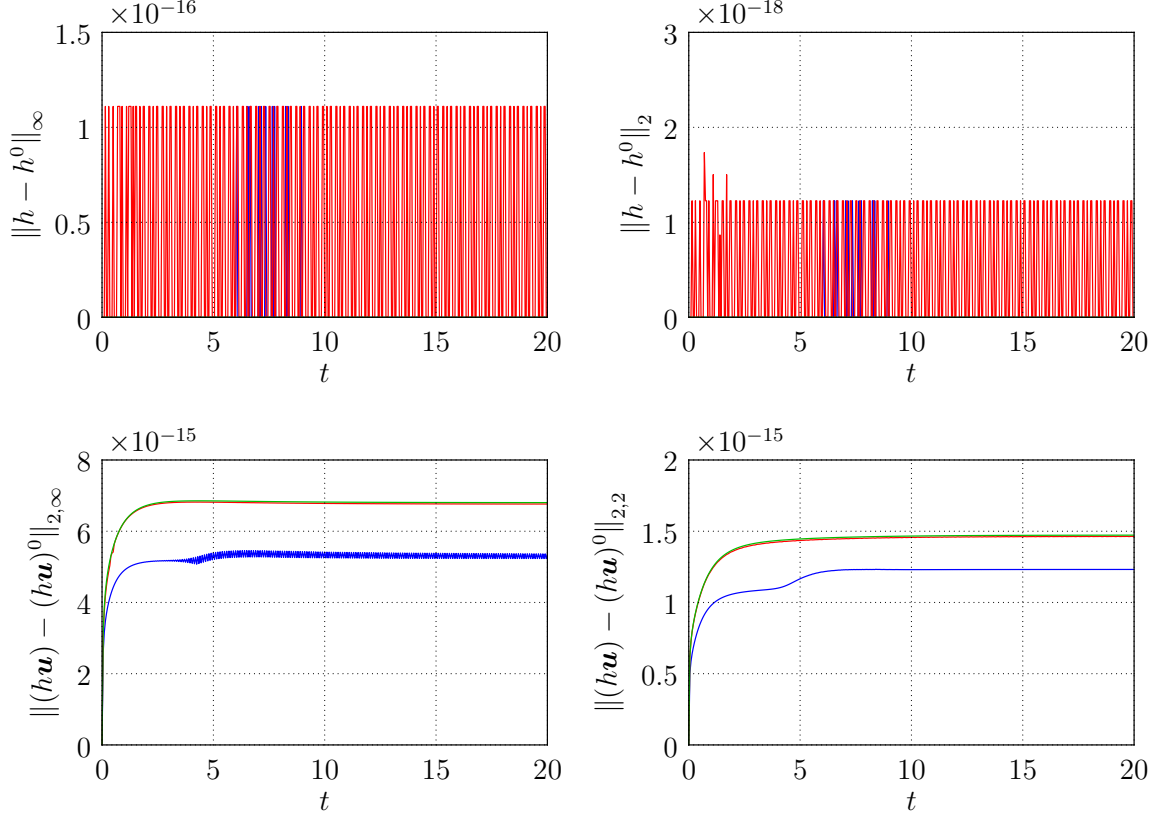
Figure 5.9: Error in height (top row) and momentum (bottom row) for the "lake at rest" test case, where the height update is computed as described in Remark 5.5. Blue: standard method, red: BDF(2) in first correction, green: BDF(2) in both corrections.

5.4.5 Flow over a hill

In the one-dimensional case with time-independent bottom topography, a steady state happens when

$$hu = C_1 \qquad \text{and} \qquad \frac{1}{2}u^2 + \frac{1}{\mathsf{Fr}^2}(h+b) = C_2 \tag{5.33}$$

with constants C_1 and C_2. This enables us to set up more complicated test cases with a non-zero velocity field, which are the extensions of the zero Froude number test case in Section 4.2.2. By combining the two equations in (5.33), a third-order equation for the height is obtained, and all positive real solutions yield physically meaningful steady states. Furthermore, other quasi-steady solutions with time-dependent bottom
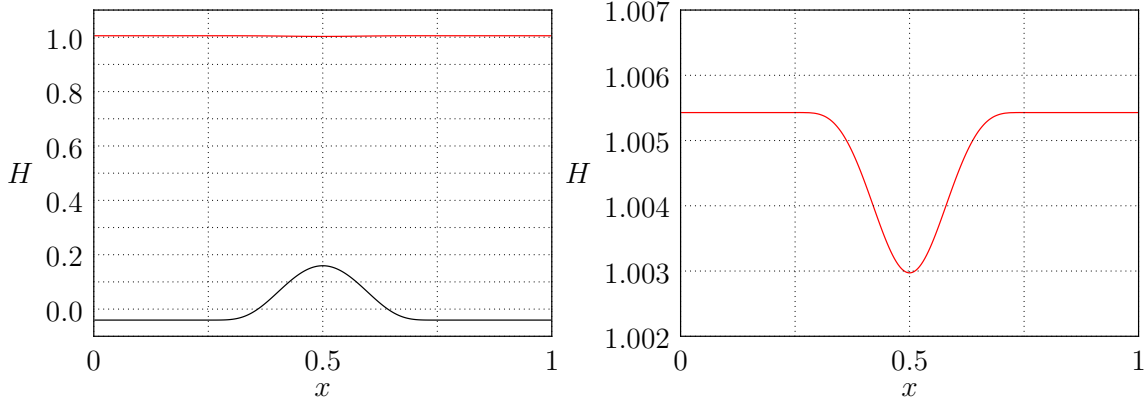
Figure 5.10: Initial total height field $H = h + b$ for the "flow over a hill" test case. Left: together with the initial bottom topography, right: detailed view of H only.

topography can be obtained by a Galilei transformation, as was already done in Section 4.2.2.

The computational domain consists of the interval $\Omega = [0, 1]$ with periodic boundary conditions. To derive the initial height, the above constants are set to $C_1 = 1$, $C_2 = 101$ and $\mathsf{Fr} = 0.1$. The bottom topography is given by

$$b(t, x) = \tilde{b}(x - v_{\mathrm{rel}} \cdot t \pmod{1}) + C = \tilde{b}(x_{\mathrm{rel}}) + C$$

with $\tilde{b}$ as given in (4.26) and $C = -\overline{\tilde{b}(x_{\mathrm{rel}})}$, to obtain a zero mean. In Figures 5.10 and 5.11, the initial fields are displayed for total height and momentum. The numerical solution is calculated for four different configurations with background velocities $v_{\mathrm{rel}} = -1.0$, -0.5, 0.0 and $+1.0$ (each corresponding to a different Galilei transformation). The tested grids have 256, 512, 1024 and 2048 grid cells, respectively, and the corresponding time steps are given by $\Delta t_{256} = 1/640$, $\Delta t_{512} = 1/1280$, $\Delta t_{1024} = 1/2560$ and $\Delta t_{2048} = 1/5120$, which is equivalent to advective Courant numbers $\mathrm{cfl}_{v_{\mathrm{rel}}=-1.0} \approx 0.08$, $\mathrm{cfl}_{v_{\mathrm{rel}}=-0.5} \approx 0.28$, $\mathrm{cfl}_{v_{\mathrm{rel}}=0.0} \approx 0.48$ and $\mathrm{cfl}_{v_{\mathrm{rel}}=1.0} \approx 0.88$ for the different choices of v_{rel}.

In addition to the "single-scale" schemes also the multiscale scheme is tested for this test case with five grid levels and blending factors $\mu_{\boldsymbol{\nu}} = (1, 1, 1, 1/2, 0)$. The schemes are integrated until the final time $t_{\mathrm{end}} = 0.1$, and the numerical solutions are compared to the exact one. In Figures 5.12 to 5.15 (Tables A.6 to A.13), the
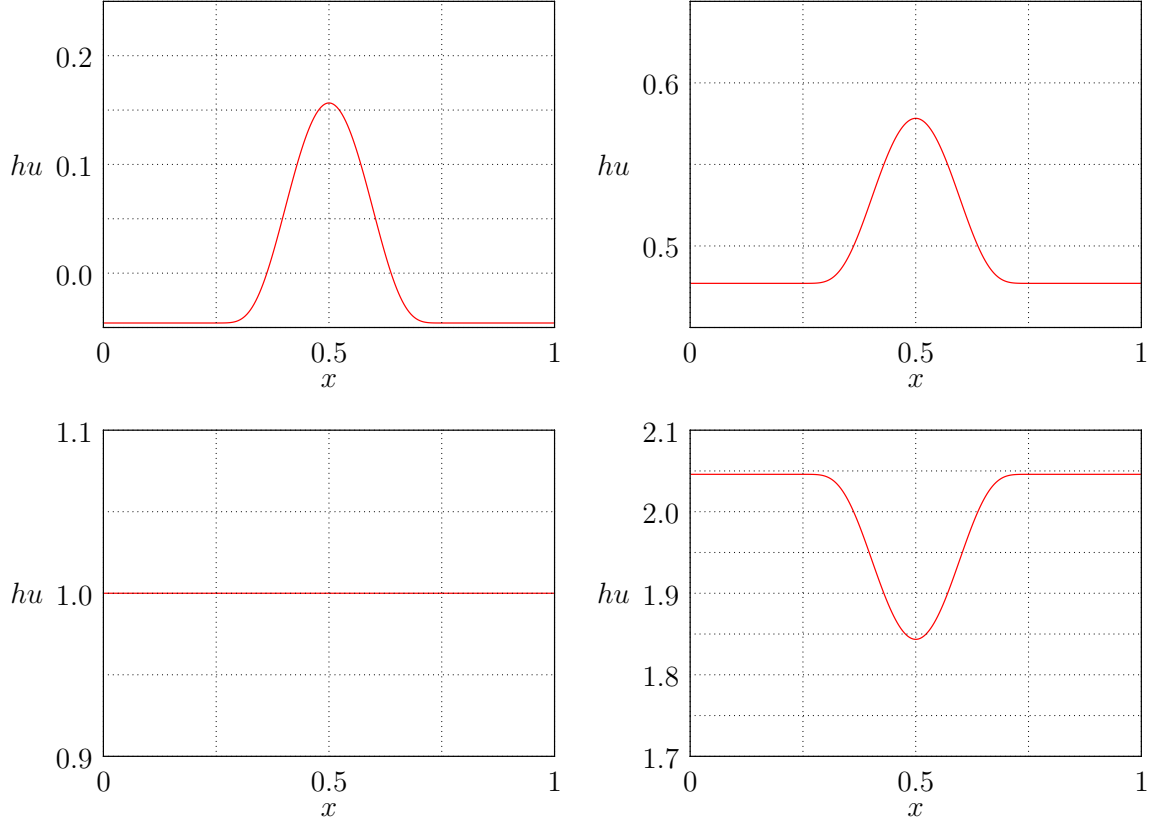
Figure 5.11: Initial momentum field for the "flow over a hill" test case and different choices of relative velocities v_{rel}. Top left: $v_{\text{rel}} = -1.0$, top right: $v_{\text{rel}} = -0.5$, bottom left: $v_{\text{rel}} = 0.0$, bottom right: $v_{\text{rel}} = +1.0$.

errors and convergence rates for the cases the different cases of v_{rel} are given. In summary the errors in the height are approximately of the same order for all different schemes, with the BDF(2)-based discretizations yielding slightly worse results. The multiscale method results in similar errors as the standard scheme. For the momentum, the standard scheme, the method with BDF(2) only in the first correction and the multiscale method yield nearly the same results. Only the method with BDF(2) in both corrections yields errors approximately twice as large. An exception is the case $v_{\text{rel}} = 0.0$. In this case, where the bottom topography is stationary, all methods result in nearly the same errors. Concerning the convergence rates, all tests again suggest second-order accuracy of the proposed schemes.
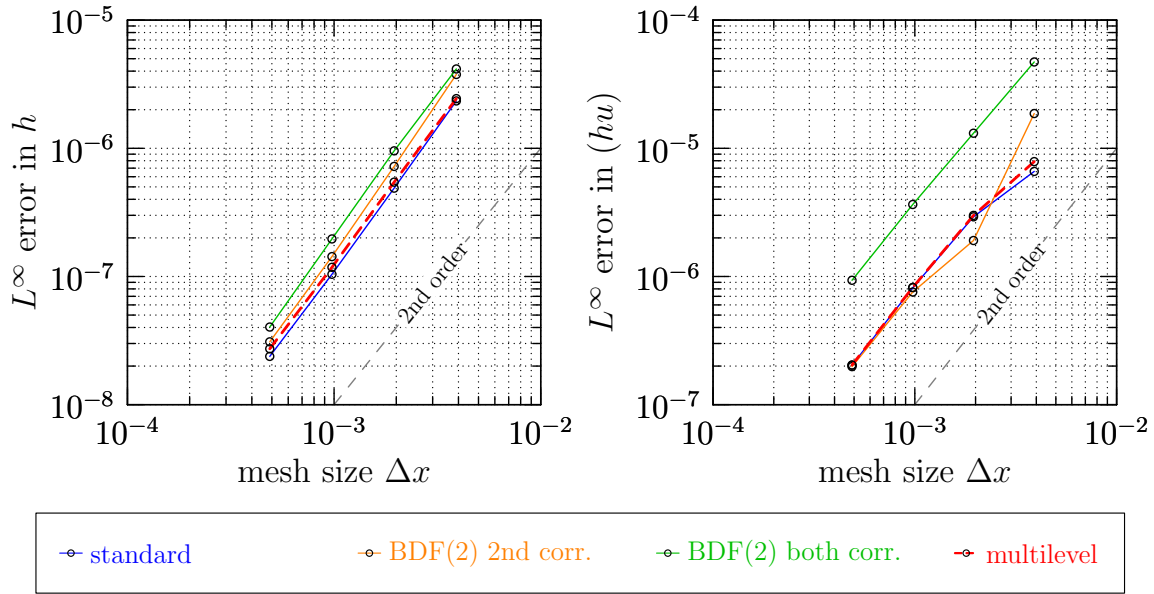
Figure 5.12: L^∞ errors in h and (hu) for the "flow over a hill" test case for $v_{\mathrm{rel}} = -1.0$ at time $t_{\mathrm{end}} = 0.1$.
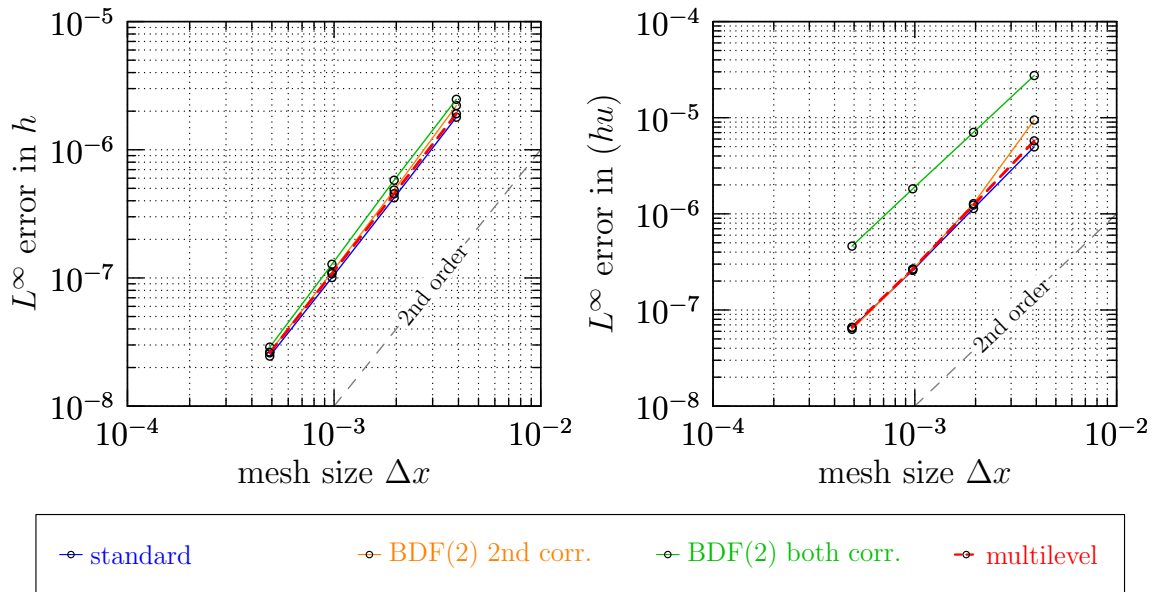
Figure 5.13: L^∞ errors in h and (hu) for the "flow over a hill" test case for $v_{\mathrm{rel}} = -0.5$ at time $t_{\mathrm{end}} = 0.1$.
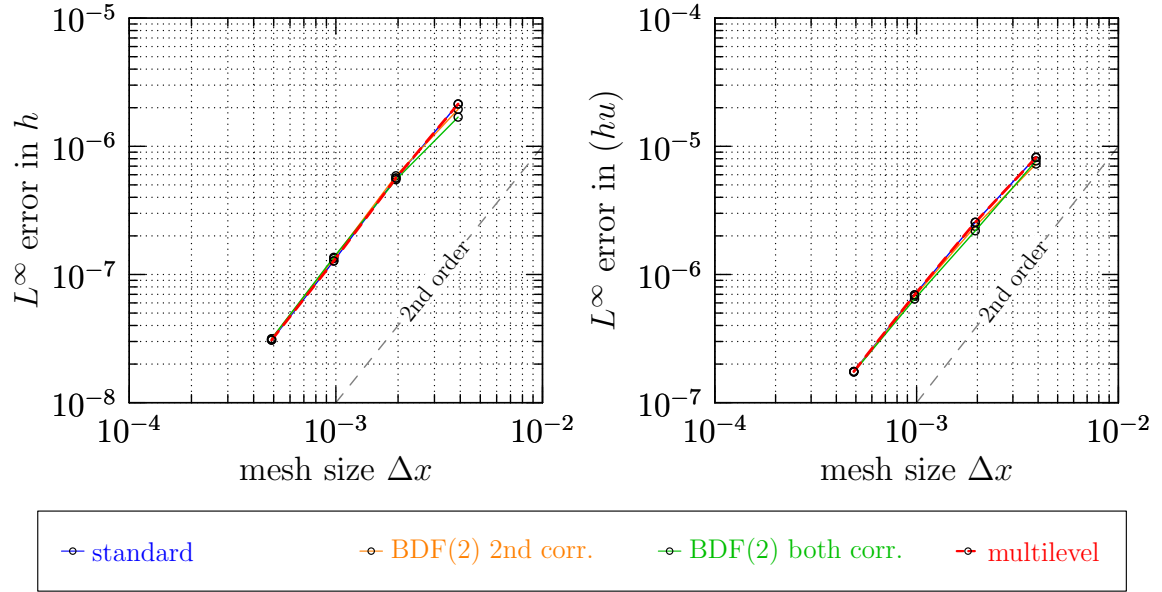
Figure 5.14: L^∞ errors in h and (hu) for the "flow over a hill" test case for $v_{\mathrm{rel}} = 0.0$ at time $t_{\mathrm{end}} = 0.1$.
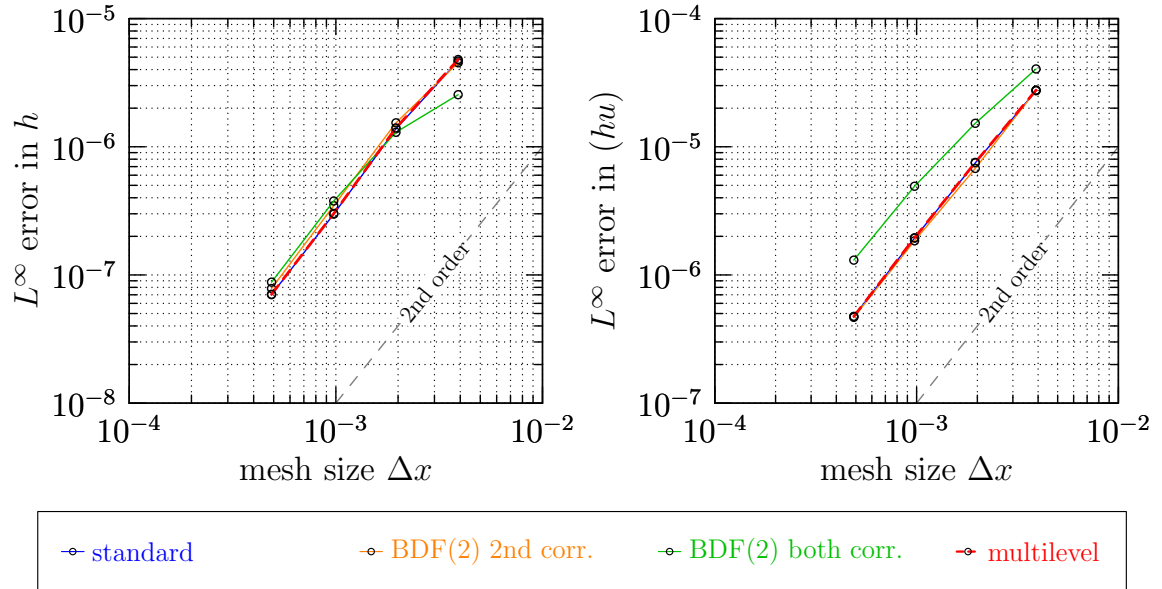
Figure 5.15: L^∞ errors in h and (hu) for the "flow over a hill" test case for $v_{\mathrm{rel}} = +1.0$ at time $t_{\mathrm{end}} = 0.1$.

5.4.6 Vortex advected over topography

As a last test case for two-dimensional flow over topography the setup from Section 4.2.3 is extended to low but non-zero Froude numbers. In the setting for the initial data a quasi one-dimensional stationary solution is superimposed with a quasi-stationary vortex, which is advected with the flow. However, as in the zero Froude number case, the vortex degenerates after some time when it interacts with the non-flat topography.

The computational domain is given by $\Omega = [0, 4] \times [0, 1]$ with periodic boundary conditions, and the Froude number is set to $\mathsf{Fr} = 0.1$. To derive the initial conditions, the bottom topography is defined by

$$b(t, x, y) = \tilde{b}(x_{\mathrm{rel}}) = \begin{cases} a & \text{if } 0 \leq x_{\mathrm{rel}} \leq 1 \\ a \cdot f(2 - x_{\mathrm{rel}}) & \text{if } 1 < x_{\mathrm{rel}} < 2 \\ 0 & \text{if } 2 \leq x_{\mathrm{rel}} \leq 3 \\ a \cdot f(x_{\mathrm{rel}} - 3) & \text{if } 3 < x_{\mathrm{rel}} < 4 \end{cases}$$

with $x_{\mathrm{rel}} = x - v_{\mathrm{rel}} \cdot t \pmod 4$ and

$$f(x) = \frac{\exp(-1/x)}{\exp(-1/x) + \exp(-1/(1-x))} \ .$$

For the computation of a stationary background state h_{bg} and u_{bg} in the x-direction, equation (5.33) is employed with $C_1 = 1$, $C_2 = 101$. The background velocity in the y-direction is given either by $v_{\mathrm{bg}} = 0$, where the vortex is advected parallel to the x-axis, or by $v_{\mathrm{bg}} = 1$, which results in a diagonal advection of the vortex. The vortex is defined as in the two-dimensional convergence test case above (Section 5.4.3). Therefore, the initial height and velocity are calculated by (5.31) and (5.32), respectively, but with the above given background states. The tangential velocity is again defined by (5.30), where the needed parameters are set to $v_{\mathrm{max}} = 1.0$ and $r_m = 0.45$. Additionally to the case of a stationary bottom topography ($v_{\mathrm{rel}} = 0$), a Galilei transformation is performed in x-direction with $v_{\mathrm{rel}} = -0.5$. The resulting initial height and vorticity fields are displayed in Figure 5.16. Furthermore, the initial profile of the bottom topography in x-direction is plotted.

The considered computational grids consist of 32×128, 64×256, 128×512 and 256×1024 grid cells. Since we already observed a more severe time step restriction for
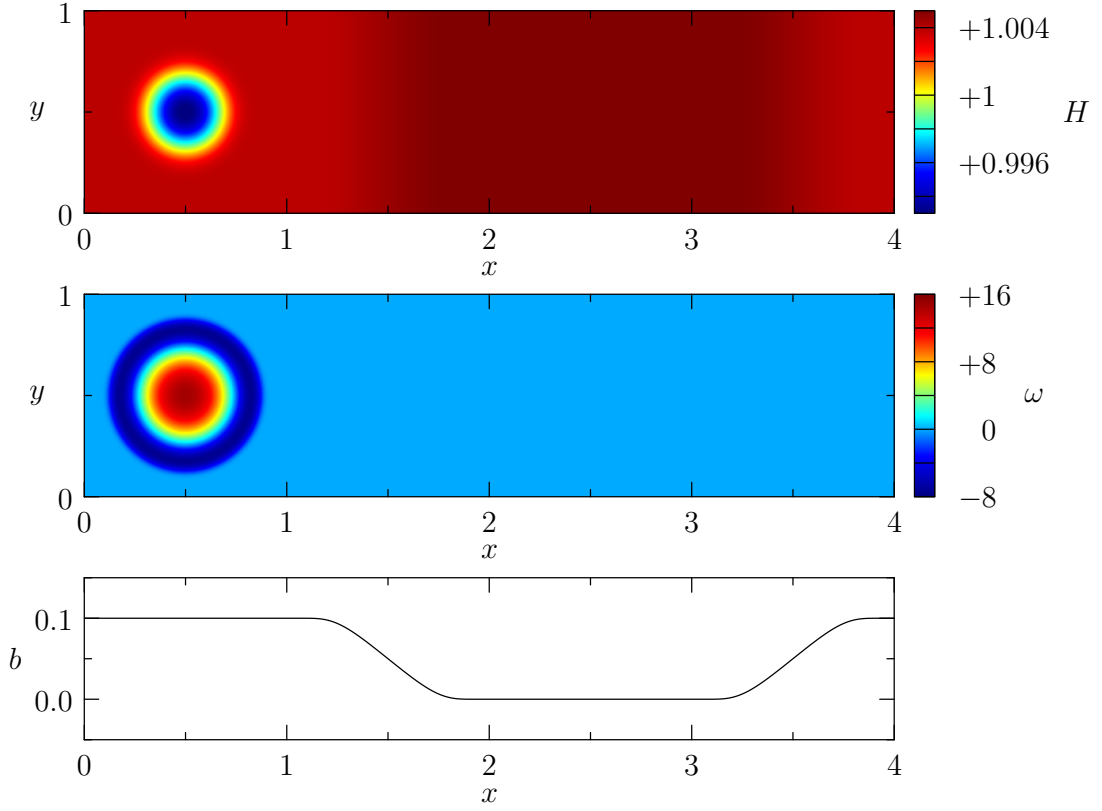
Figure 5.16: Initial height (top) and vorticity (middle) distribution and variation of the bottom topography in x-direction (bottom) for the "vortex over topography" test case.

the method with BDF(2) discretizations in both corrections, different time step sizes are chosen for this scheme compared to the other two. For $v_{\text{rel}} = 0$ they are given by $\Delta t_{32} = 1/160$, $\Delta t_{64} = 1/320$, $\Delta t_{128} = 1/640$ and $\Delta t_{256} = 1/1280$ for the standard method and the scheme with BDF(2) in the second correction. For the method with BDF(2) in both corrections, the time steps are set to $\Delta t_{32} = 1/250$, $\Delta t_{64} = 1/500$, $\Delta t_{128} = 1/1000$ and $\Delta t_{256} = 1/2000$. This corresponds (for $v_{\text{bg}} = 1$) to advective Courant numbers $\text{cfl}_{\text{adv}} \approx 0.72$ and $\text{cfl}_{\text{adv}} \approx 0.46$, respectively. For $v_{\text{rel}} = -0.5$, the time steps are chosen to be either $\Delta t_{32} = 1/125$, $\Delta t_{64} = 1/250$, $\Delta t_{128} = 1/500$ and $\Delta t_{256} = 1/1000$ (standard method, BDF(2) in second correction), or $\Delta t_{32} = 1/200$, $\Delta t_{64} = 1/400$, $\Delta t_{128} = 1/800$ and $\Delta t_{256} = 1/1600$ (BDF(2) in both corrections). This results in advective Courant numbers $\text{cfl}_{\text{adv}} \approx 0.8$ and $\text{cfl}_{\text{adv}} \approx 0.5$ (for $v_{\text{bg}} = 1$).

Table 5.1: Experimental convergence rates in h for the "vortex over topography" test case for $v_{\mathrm{rel}} = 0.0$ at time $t = 2$.

method	norm	rate γ_{32}^{64}	rate γ_{64}^{128}	rate γ_{128}^{256}
standard scheme	L^2	0.721	1.359	1.962
	L^∞	0.241	1.476	1.985
BDF(2) only in 2nd correction	L^2	0.698	1.358	1.962
	L^∞	0.209	1.487	1.972
BDF(2) in both corrections	L^2	0.733	1.354	1.963
	L^∞	0.246	1.432	2.004

Table 5.2: Experimental convergence rates in $(h\boldsymbol{u})$ for the "vortex over topography" test case for $v_{\mathrm{rel}} = 0.0$ at time $t = 2$.

method	norm	rate γ_{32}^{64}	rate γ_{64}^{128}	rate γ_{128}^{256}
standard scheme	L^2	0.857	1.421	1.851
	L^∞	0.687	1.200	1.792
BDF(2) only in 2nd correction	L^2	0.857	1.421	1.851
	L^∞	0.686	1.199	1.792
BDF(2) in both corrections	L^2	0.910	1.408	1.847
	L^∞	0.765	1.160	1.811

The computations are performed until the final time $t_{\mathrm{end}} = 3$. At this time, the center of the vortex is located at the upward slope of the bottom topography ($x \approx 3.5$ for $v_{\mathrm{rel}} = 0$). In Figure 5.17 the resulting height and vorticity fields are displayed using the standard scheme on a grid with 256×1024 grid cells with $v_{\mathrm{rel}} = 0$ and $v_{\mathrm{bg}} = 1$. Since the results from the schemes with BDF(2)-based discretizations are nearly indistinguishable from this one, they are not separately plotted. Additionally, details of the solution are given for different grids at $t = 2$ in Figure 5.18. As in the zero Froude number case, the results suggest convergence to the unknown solution. The corresponding plots for $v_{\mathrm{rel}} = -0.5$ and $v_{\mathrm{bg}} = 1$ are given in Figures 5.19 and 5.20. Apart from the Galilei transformation the obtained solutions are almost the same as for the case $v_{\mathrm{rel}} = 0$, also in this case. Only for coarse grids minor differences are visible.

Similar to the zero Froude number case the experimental order of convergence is computed for the different single-scale methods by comparing the numerical solutions on different grids. The results for the height field are given in Tables 5.1 and 5.3
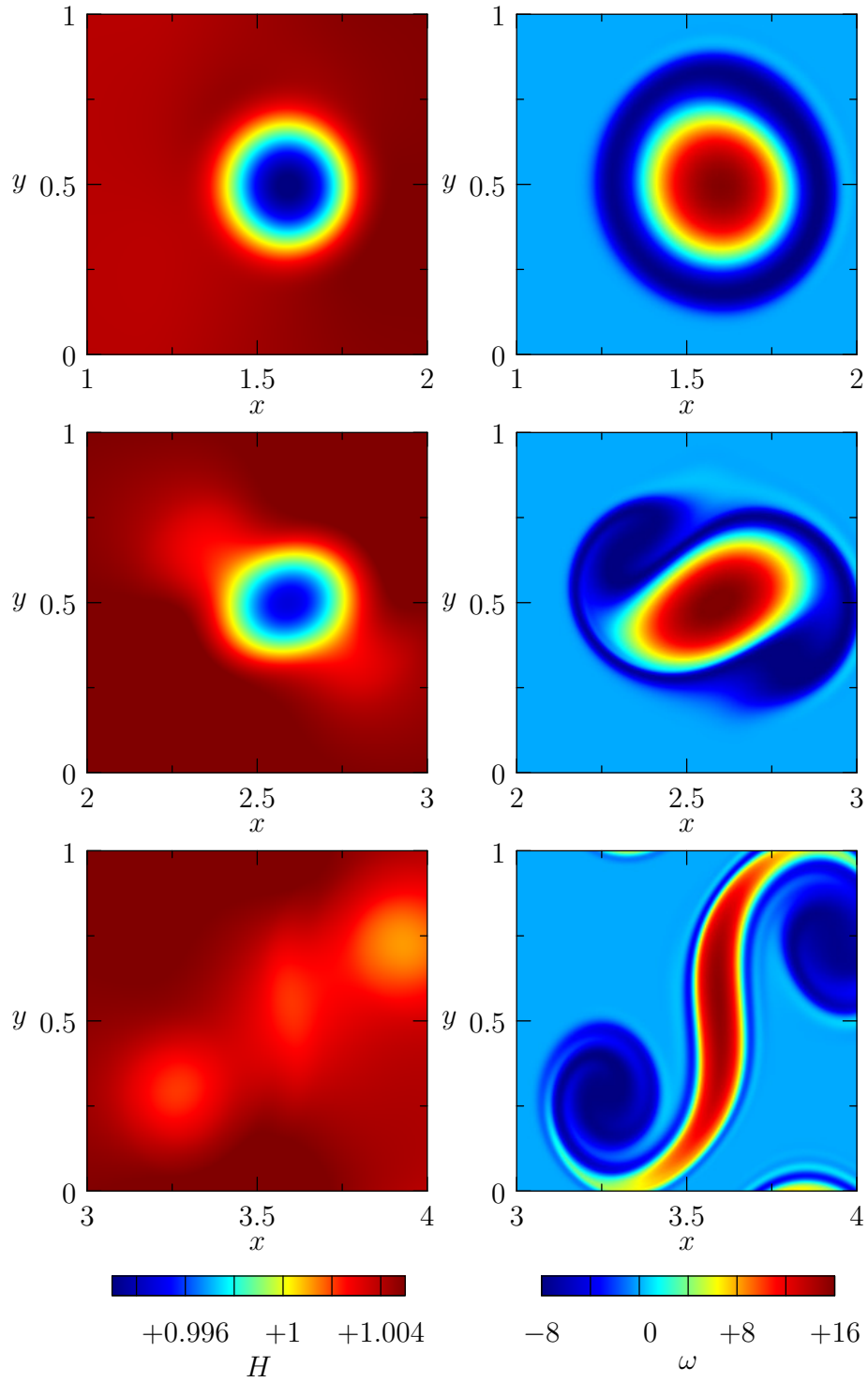
Figure 5.17: Height (left) and vorticity (right) fields for the "vortex over topography" test case with $v_{\mathrm{rel}} = 0$ and 256×1024 grid cells. Details of the computational domain at times $t = 1$ (top), $t = 2$ (middle) and $t_{\mathrm{end}} = 3$ (bottom).
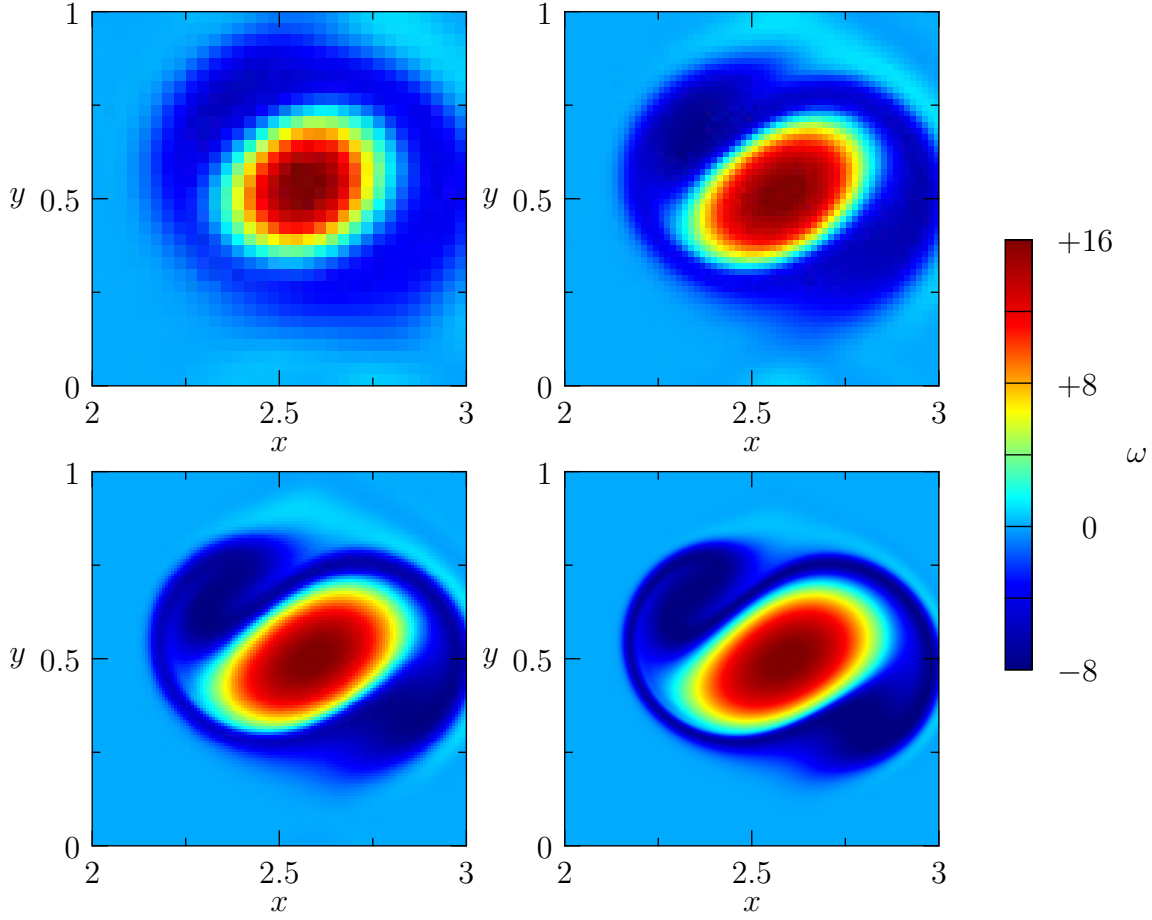
Figure 5.18: Vorticity distribution for the "vortex over topography" test case with $v_{\text{rel}} = 0$ at $t = 2$. Details of the computational domain using 32×128 (top left), 64×256 (top right), 128×512 (bottom left) and 256×1024 (bottom right) grid cells.

for $v_{\text{rel}} = 0$ and $v_{\text{rel}} = -0.5$, respectively. For the momentum, they are given in Tables 5.2 and 5.4. For coarse grids the convergence rates are very low in most cases. However, as the grid resolution is increased, the rates clearly tend towards second-order convergence.

There are still some problems in the simulation of low Froude number shallow water flows with the presented schemes, which are due to the solution method for the auxiliary system in the predictor step. To show this, some of the above tests are repeated with background velocity $v_{\text{bg}} = 0$, which means that the background flow is parallel to the x-axis. For the computation of the presented numerical results
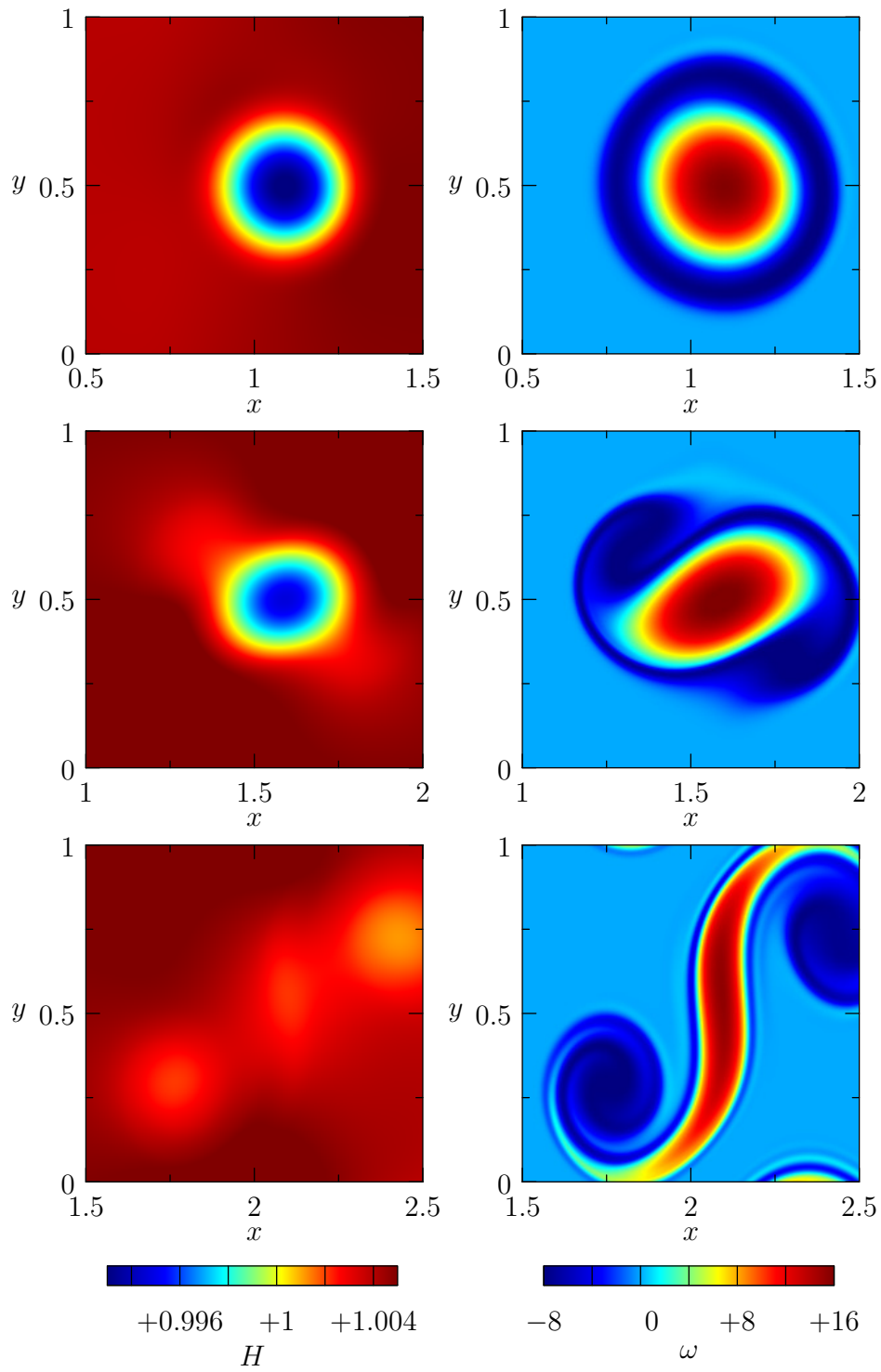
Figure 5.19: Same as Figure 5.17, but with $v_{\mathrm{rel}} = -0.5$.
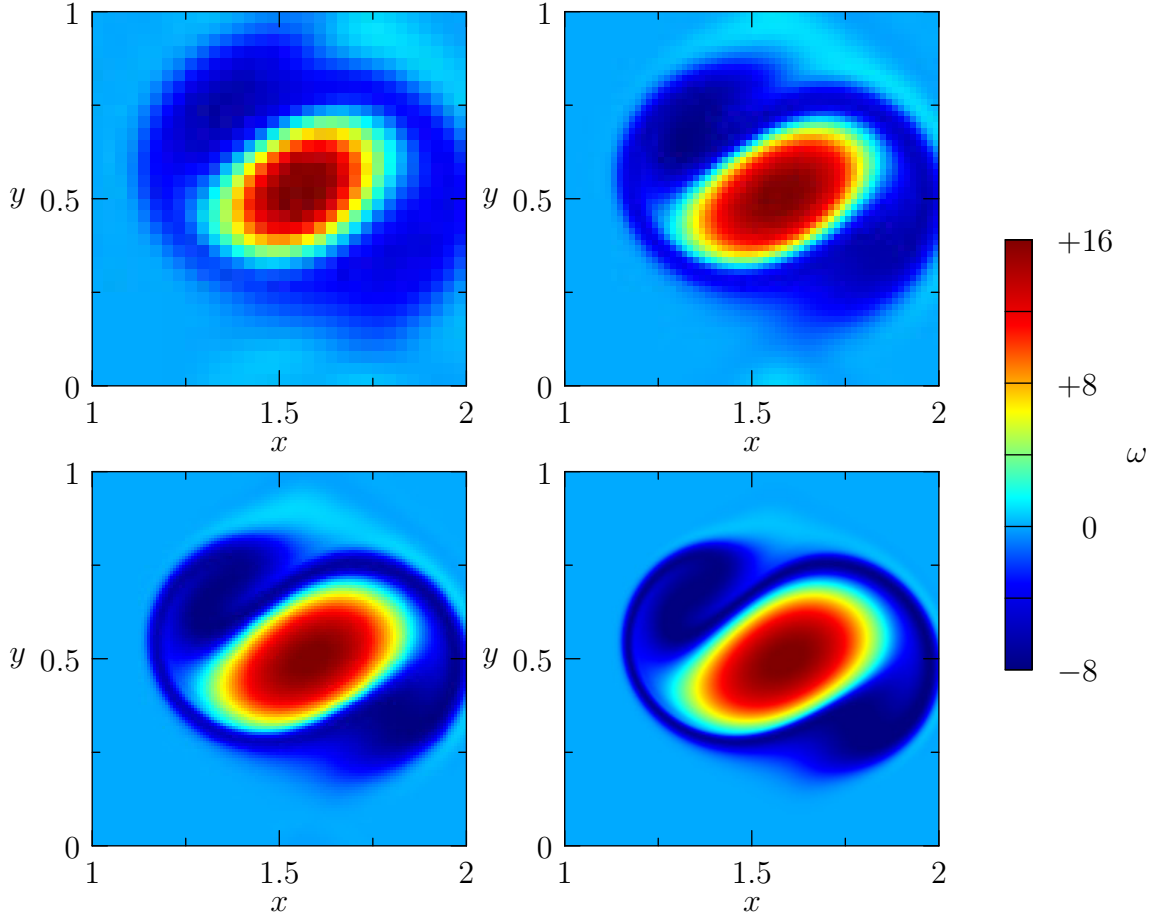
Figure 5.20: Same as Figure 5.18, but with $v_{\mathrm{rel}} = -0.5$.

the standard scheme is used, but the same findings are true for the BDF(2)-based discretizations. In Figure 5.21 the height field is displayed at time $t = 2$ computed on a grid with 32×128 cells. The top image shows the solution for $v_{\mathrm{rel}} = 0$, while the bottom one shows the result for $v_{\mathrm{rel}} = -1$. With these configurations the appearance of numerical artifacts is observed: for the case $v_{\mathrm{rel}} = 0$, stripes in the x-direction appear in areas outside the vortex. These effects are amplified for the case $v_{\mathrm{rel}} = -1$, in which the vortex is essentially stationary. They appear as oscillations in both coordinate directions adjacent to the vortex. Moreover, these artifacts do not disappear for refined grids (not shown).

For comparison, the same test case is computed with a modified scheme, in which the pressureless equations are replaced by the equations (5.4) with $\Phi = 0.5$ in the

Table 5.3: Experimental convergence rates in h for the "vortex over topography" test case for $v_{\mathrm{rel}} = -0.5$ at time $t = 2$.

method	norm	rate γ_{32}^{64}	rate γ_{64}^{128}	rate γ_{128}^{256}
standard scheme	L^2	0.926	1.261	1.930
	L^∞	1.002	1.400	1.807
BDF(2) only in 2nd correction	L^2	0.921	1.250	1.928
	L^∞	1.026	1.373	1.804
BDF(2) in both corrections	L^2	1.398	1.847	1.953
	L^∞	1.703	1.958	1.919

Table 5.4: Experimental convergence rates in $(h\boldsymbol{u})$ for the "vortex over topography" test case for $v_{\mathrm{rel}} = -0.5$ at time $t = 2$.

method	norm	rate γ_{32}^{64}	rate γ_{64}^{128}	rate γ_{128}^{256}
standard scheme	L^2	1.000	1.289	1.838
	L^∞	0.692	1.087	1.750
BDF(2) only in 2nd correction	L^2	1.000	1.289	1.838
	L^∞	0.690	1.087	1.750
BDF(2) in both corrections	L^2	1.314	1.677	1.914
	L^∞	1.138	1.645	1.954

predictor step. These are essentially the shallow water equations with Froude number $\sqrt{0.5}$. Note, that this modification also leads to some changes in the correction steps. Furthermore, a well-balanced discretization is applied to account for the term $\Phi h \nabla (h + b)$. For the computation of the numerical fluxes of the auxiliary system the HLLE-Riemann solver (Einfeldt, 1988) is used. The results from this method are given in Figure (5.22). Clearly, the unphysical oscillations have vanished. This suggests that the solution of the pressureless equations using an exact Riemann solver introduces not enough numerical viscosity for height variations, which are advected parallel to the interface of a cell. On the other hand, the HLLE solver introduces viscosity for such waves, and small numerical errors cannot accumulate. However, the scope of this thesis is on the development of a multiscale method. While the precise implementation of the auxiliary system should be further considered in future work, it is not so important at this stage.
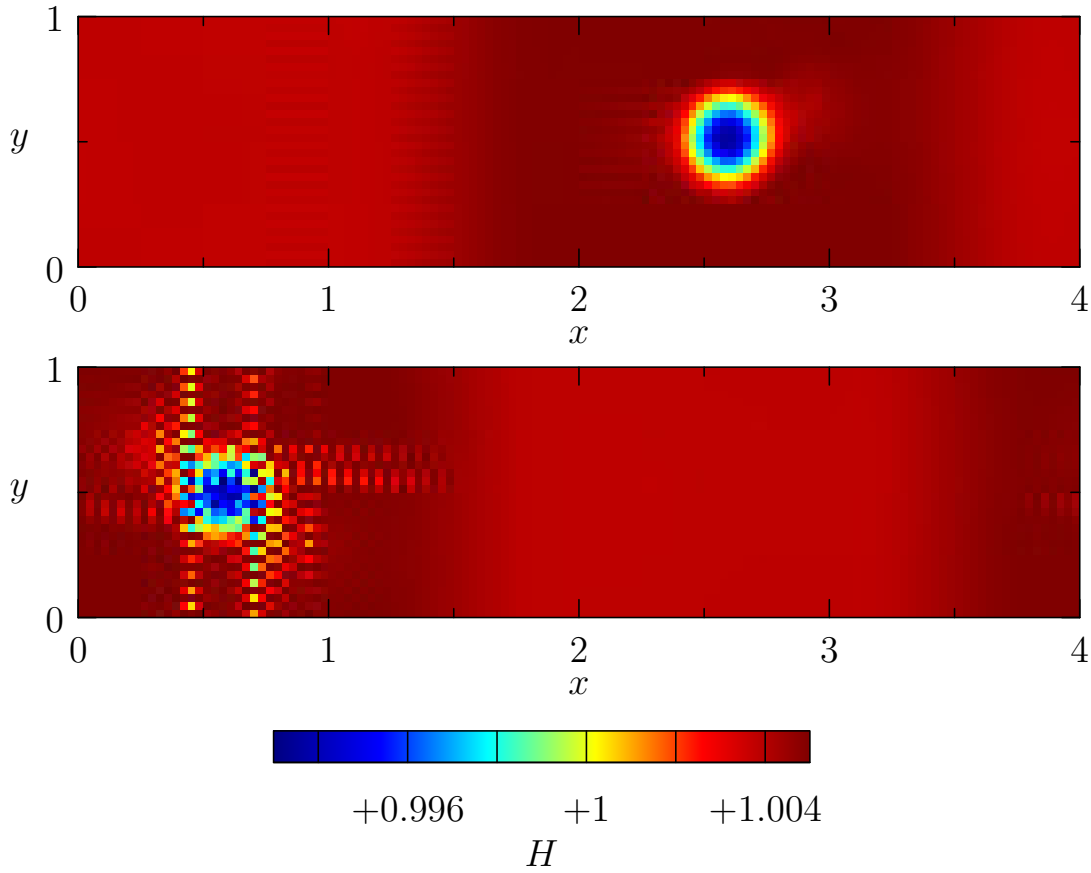
Figure 5.21: Numerical solution (height field) of the "vortex over topography" test case using the pressureless equations as predictor at time $t = 2$ on a grid with 32×128 grid cells. Vortex advected parallel to the x-axis ($v_{\mathrm{bg}} = 0$). Top: $v_{\mathrm{rel}} = 0$, bottom: $v_{\mathrm{rel}} = -1$.

5.4.7 Balanced modes in presence of time dependent bottom topography

In a final test case, the proposed schemes are tested for their ability to relax to non-trivial balanced states in the presence of bottom topography varying in time. In order to do so, the test case from Section 3.3.2 is extended to the fully nonlinear shallow water equations. This test case is done in one space dimension and the computational domain is given by $\Omega = [0, 100]$.
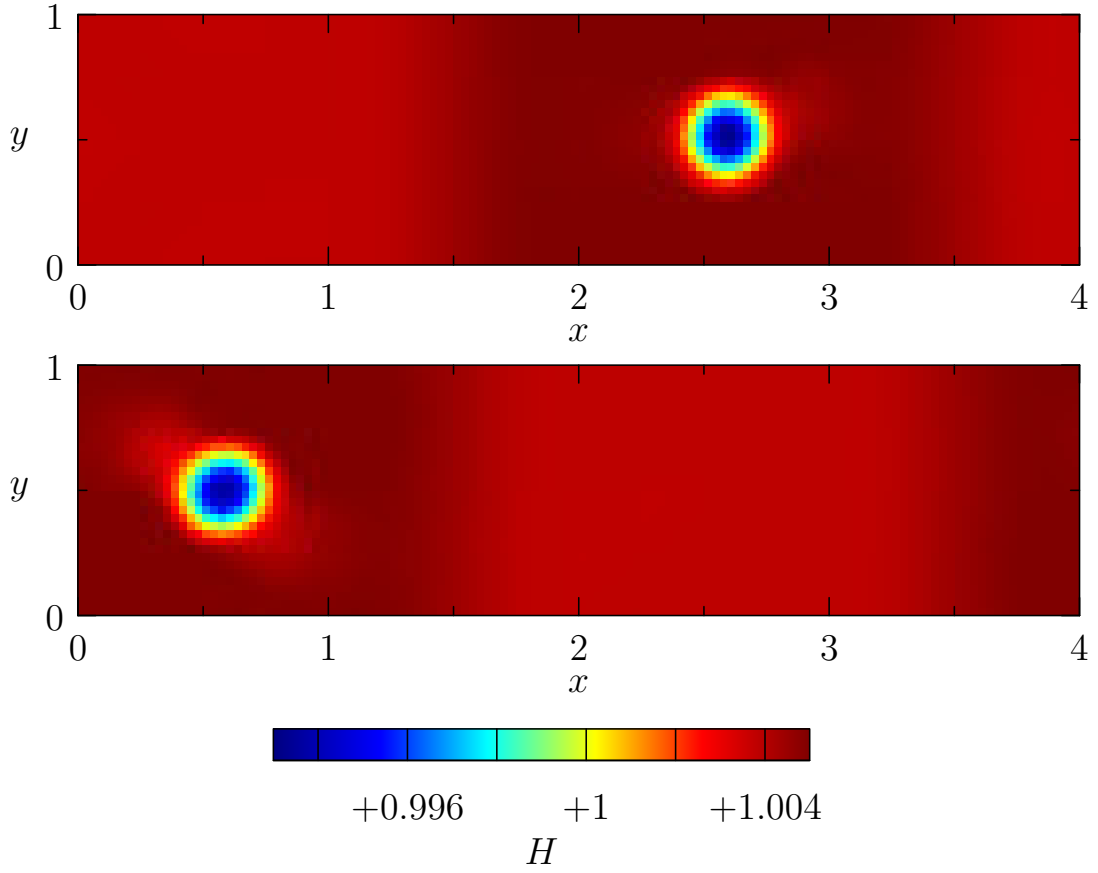
Figure 5.22: Same as Figure 5.21, but using the equations (5.4) with $\Phi = 0.5$ as predictor.

The bottom topography is given by

$$b(t, x) = \mathsf{Fr} \sin (\omega\, t)\, \tilde{q}(x - x_0)\ ,$$

where $\tilde{q}$ is defined by (3.19) with the same parameters as in Section 3.3.2. Initially the fluid is at rest $(u \equiv 0)$ with height $h \equiv 1$. When the flow is in balance, the findings from Chapter 2 imply that for small Froude numbers the perturbations in height and momentum should be also small, and the dynamics primarily happen in the linear regime. This means that the solution should be governed by the asymptotic solution obtained for the linearized shallow water equations. Translated to the given initial value problem and bottom topography, the total height perturbation and velocity

should be governed by

$$H_{\mathrm{asy}}(t,x) - H_0 = -\frac{\mathsf{Fr}^3}{H_0}\omega \sin\left(\omega\, t\right)\tilde{h}(x-x_0)$$

with $\tilde{h}$ as given in (3.21) and

$$u_{\mathrm{asy}}(t,x) = \frac{\mathsf{Fr}}{H_0}\cos\left(\omega\, t\right)\tilde{u}(x-x_0)\ ,$$

where $\tilde{u}$ is defined by (3.23).

In the presented computations, the Froude number is set to $\mathsf{Fr} = 0.01$ and the total background height is $H_0 = 1$. The computational grid has 256 grid cells, and the fixed time step is given by $\Delta t = 0.3125$, which corresponds to an advective Courant number $\mathrm{cfl}_{\mathrm{adv}} \approx 0.008$ when the flow is essentially balanced. The Courant number corresponding to the transport of gravity waves, on the other hand, is $\mathrm{cfl}_{\mathrm{grav}} \approx 80$.

For this test case the BDF(2)-based schemes are not initialized with an initial step by the standard scheme. Instead, the required state at $t^{-1} = -0.3125$ is set to the balanced solution with flat bottom topography. However, compared to an initialization using the standard scheme the findings are qualitatively the same. For the multiscale method five grid levels are used with a scale-dependent blending given by $\mu_{\boldsymbol{\nu}} = (2/3, 1/2, 1/3, 1/6, 0)$.

Given the above initial conditions for $t = 0$, the height is in balance with the initial bottom topography. However, the temporal change of the latter introduces divergence into the velocity field, which, in turn, results in higher-order height perturbations. In Figure 5.23, the numerical results are displayed together with the asymptotic solution for the first six time steps using the standard method. Both, the height perturbations and the momentum field computed with the numerical scheme oscillate around the balanced state, but they do not relax to it. Additionally, the numerically calculated height perturbations are about 8 times larger compared to the asymptotic solution. The BDF(2)-based discretizations result in a completely different behavior, which is displayed in Figures 5.24 and 5.25. Here, the initial deviations from the balanced state vanish after only a few time steps. After the fourth time step the numerical solution is nearly indistinguishable from the asymptotic solution. This behavior is also reproduced by the multiscale method, for which the results are given in Figure

5.26. These results are in good agreement with the findings from Section 3.3.2 in case of the linearized shallow water equations.

In a second run, the simulation is started at $t = 0.15$, and the bottom topography is assumed to be flat before this time. At this time, when the bottom topography switches instantaneously to another state, both, height and momentum are not in balance. This leads to much bigger initial deviations from the asymptotic solution, as can be seen in Figure 5.27 for the standard scheme (note the different scaling in the y-axis for the height perturbation). To evaluate the long-term behavior, the numerical solution is additionally plotted after time steps 19 to 24 in Figure 5.28. Also in this case the solution of the standard method does not relax to the balanced state, but rather oscillates around it. Only the long-wave perturbations are diminished with time. Here, the height perturbations computed by the numerical scheme are about two orders of magnitude larger than those predicted by the asymptotic solution. For the momentum, the amplitude of the numerical solution is also about three times larger than the predicted balanced state.

The corresponding results for the BDF(2)-based methods are displayed in Figures 5.29, 5.30 (BDF(2) only in first correction) and Figures 5.31, 5.32 (BDF(2) in both corrections). The behavior of the schemes is the same as in the first setup. After initial deviations, which are of the same order as for the standard method, the numerical solutions essentially relax to the balanced state predicted by the asymptotic solution. Only in the height field, very high-wave-number small-amplitude deviations persist. Additional tests suggest that these artifacts are due to the fact that the explicit predictor cannot cope with too high-wave-number modes at these large Courant numbers. In this part of the scheme, a two-stage Runge-Kutta method is used for the time discretization. Since the gravity waves are generated by the "source term" of the predictor, which is always evaluated at the old time level, high-wave-number gravity waves get very much distorted in the second stage of the Runge-Kutta scheme. This can eventually lead to instabilities, if these parts of the solution become to large.

The results of the multiscale method are given in Figures 5.33 and 5.34. Qualitatively, the result is similar to that obtained with the scheme which has a BDF(2)-based second correction. However, the scale dependent blending of the two methods leads to even larger very high-wave-number deviations, whose amplitude is of the order of the height perturbations.

Figure 5.23: Numerical solution of the "multiscale test case" after the first six time steps using the standard scheme (black line) on a grid with 256 cells, Fr $= 0.01$. Left column: total height perturbation, middle column: momentum, right column: bottom topography. Each step n is one row. Asymptotic solution (from linear case) is plotted as dashed lines.
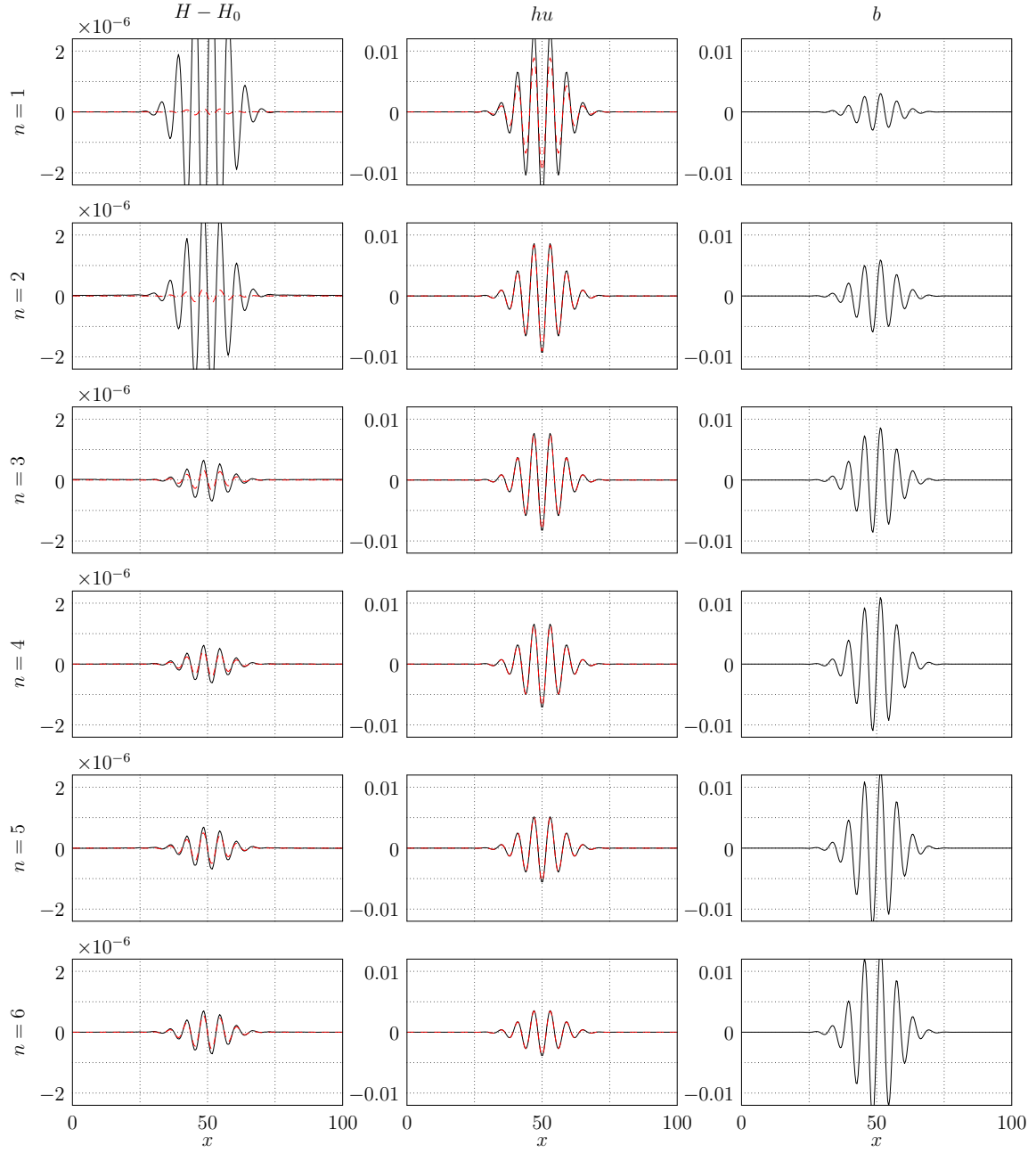
Figure 5.24: Same as Figure 5.23, but using the scheme with BDF(2) only in the second correction.
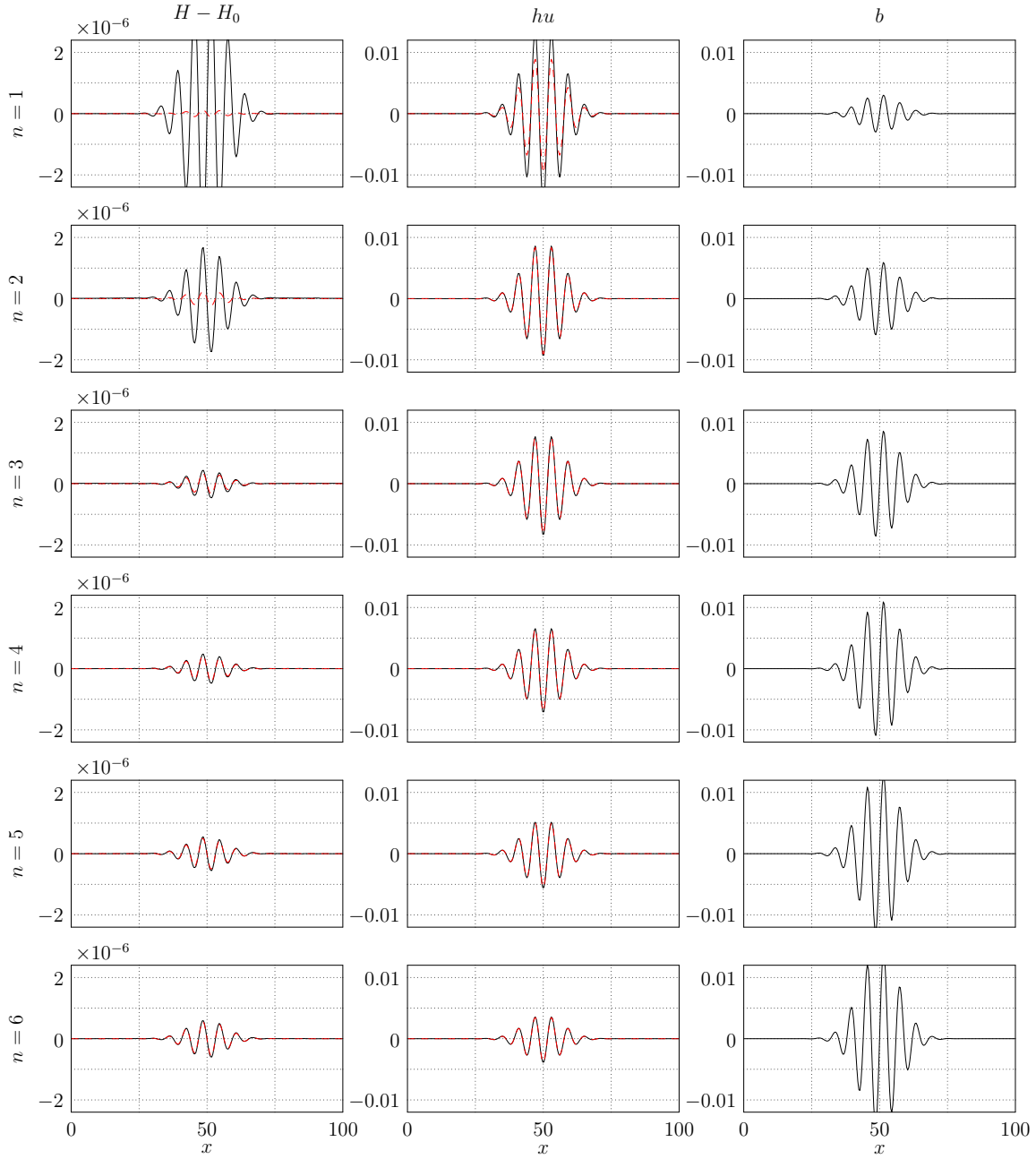
Figure 5.25: Same as Figure 5.23, but using the scheme with BDF(2) in both corrections.
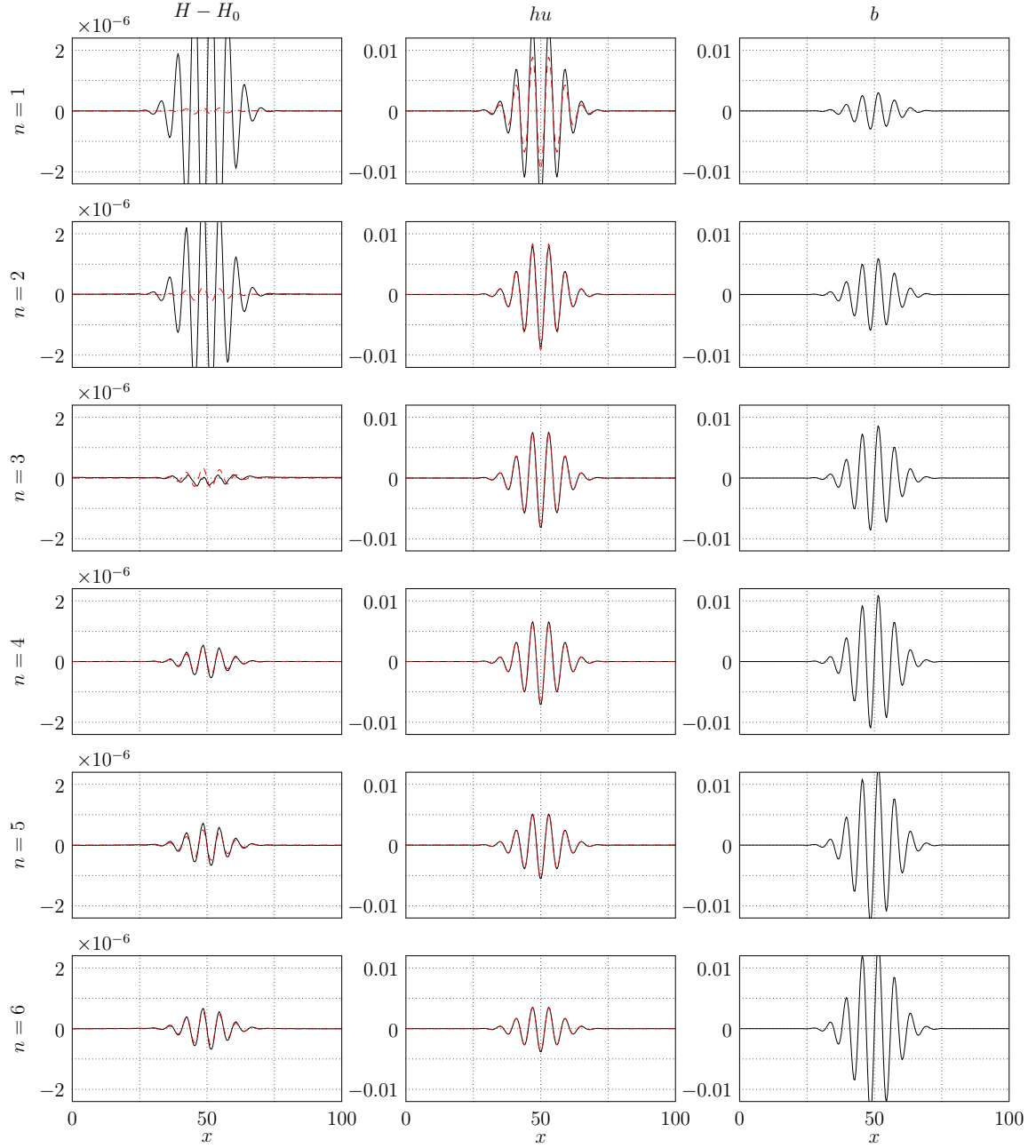
Figure 5.26: Same as Figure 5.23, but using the multiscale scheme.
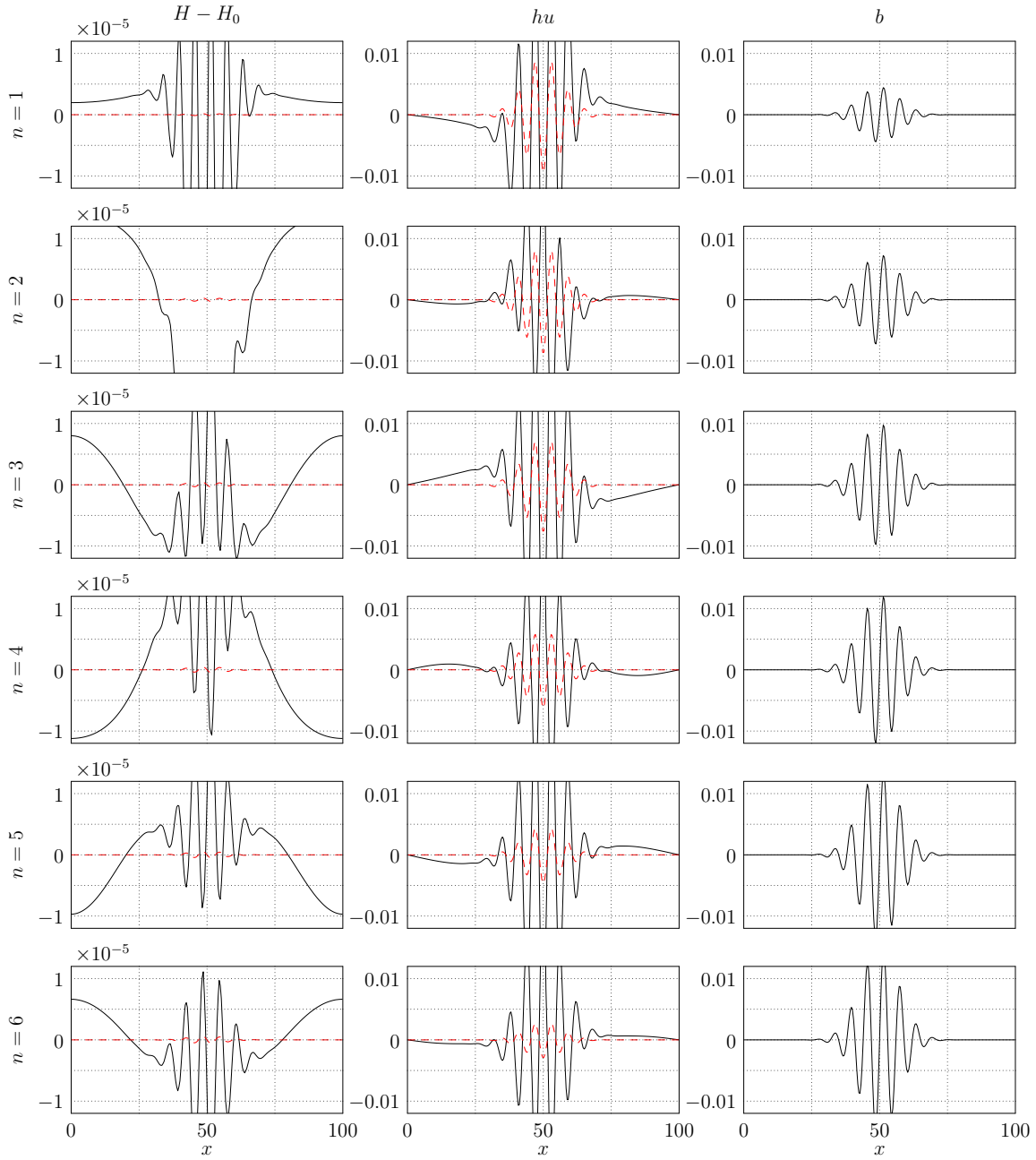
Figure 5.27: Completely unbalanced initial data. Numerical solution of the "multiscale test case" after the first six time steps using the standard scheme (black line) on a grid with 256 cells, Fr = 0.01. Asymptotic solution (from linear case) is plotted as dashed lines.
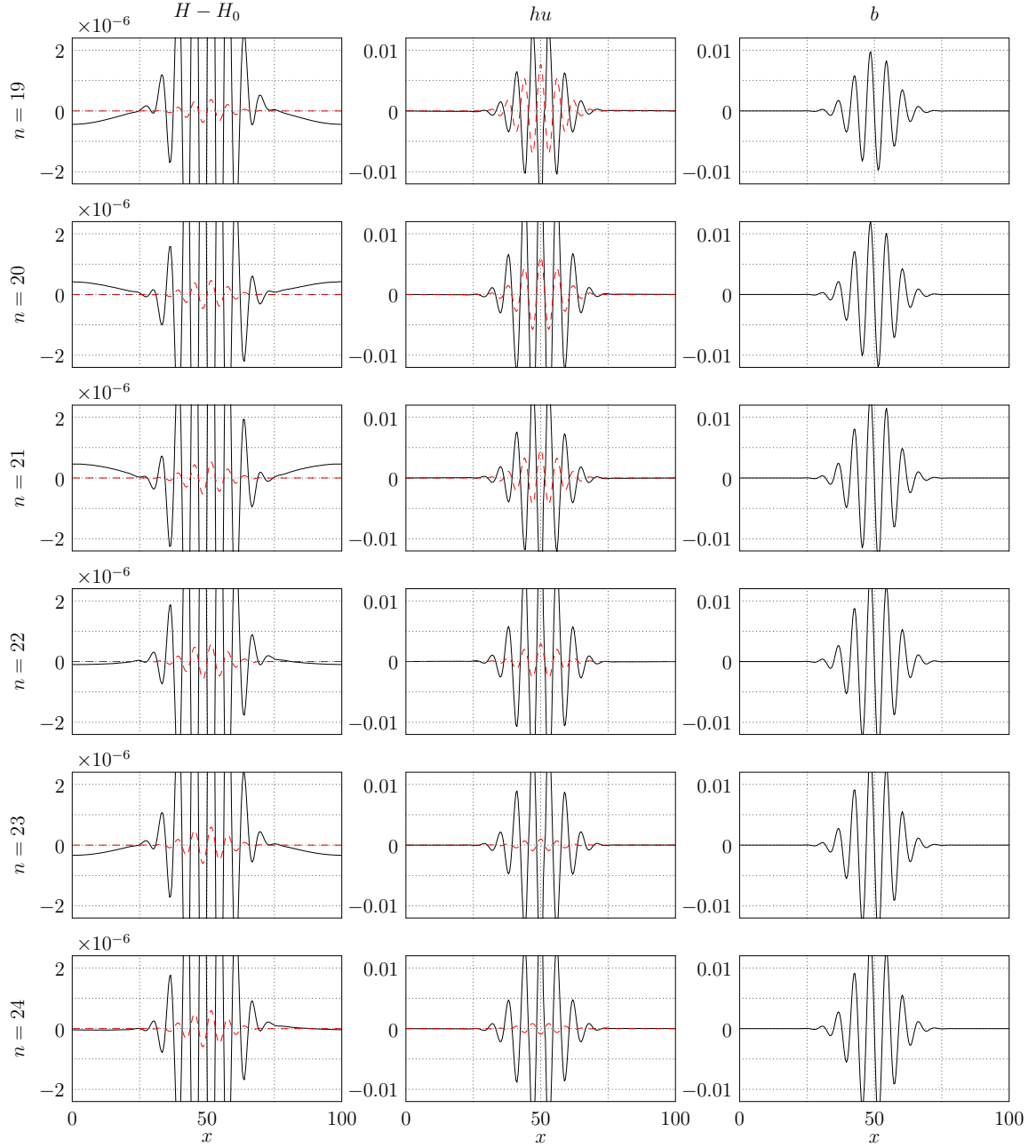
Figure 5.28: Same as Figure 5.27 with numerical solution after time steps 19 to 24.
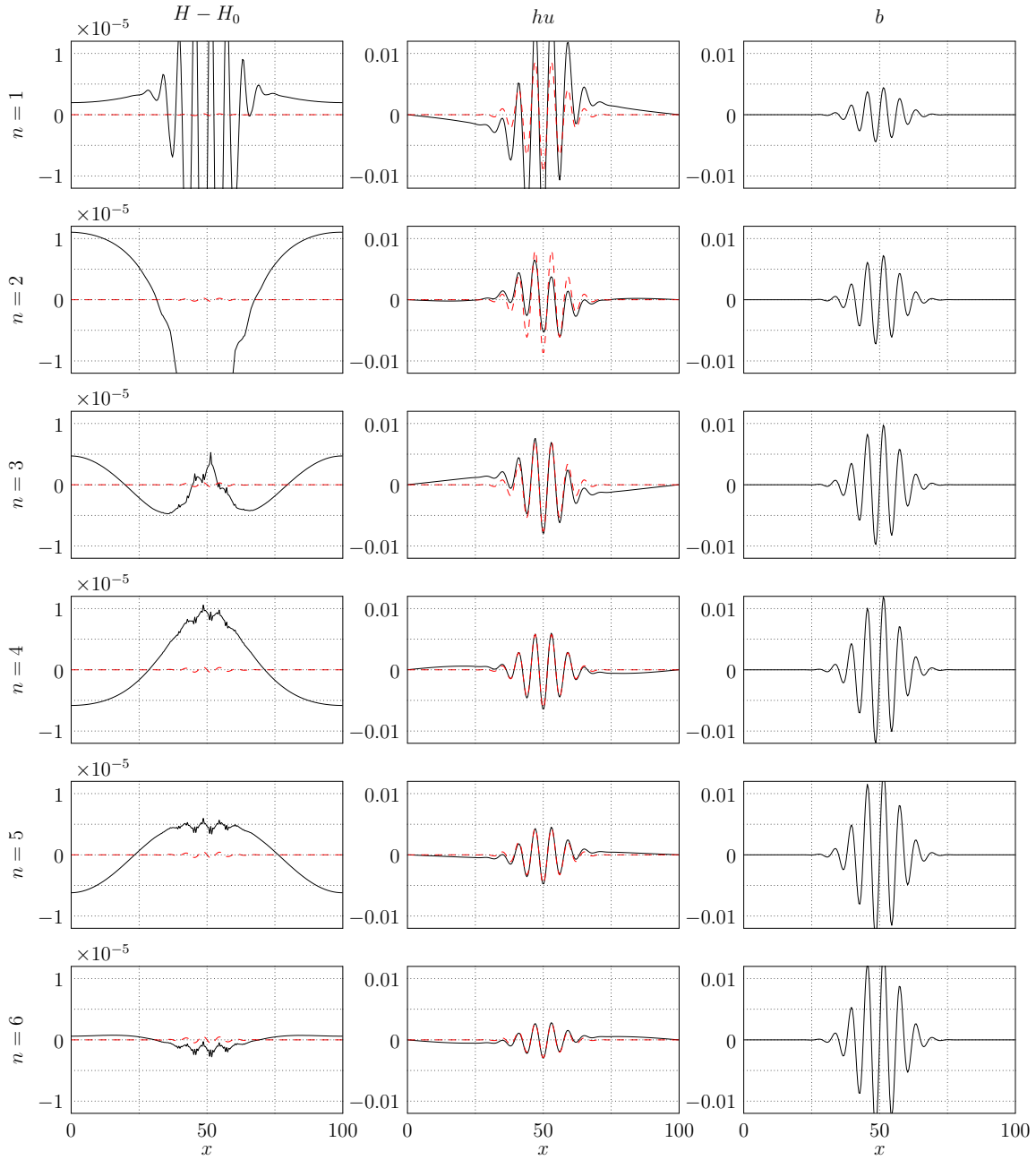
Figure 5.29: Completely unbalanced initial data. Numerical solution of the "multiscale test case" after the first six time steps using the scheme with BDF(2) only in the second correction (black line) on a grid with 256 cells, Fr $= 0.01$. Asymptotic solution (from linear case) is plotted as dashed lines.
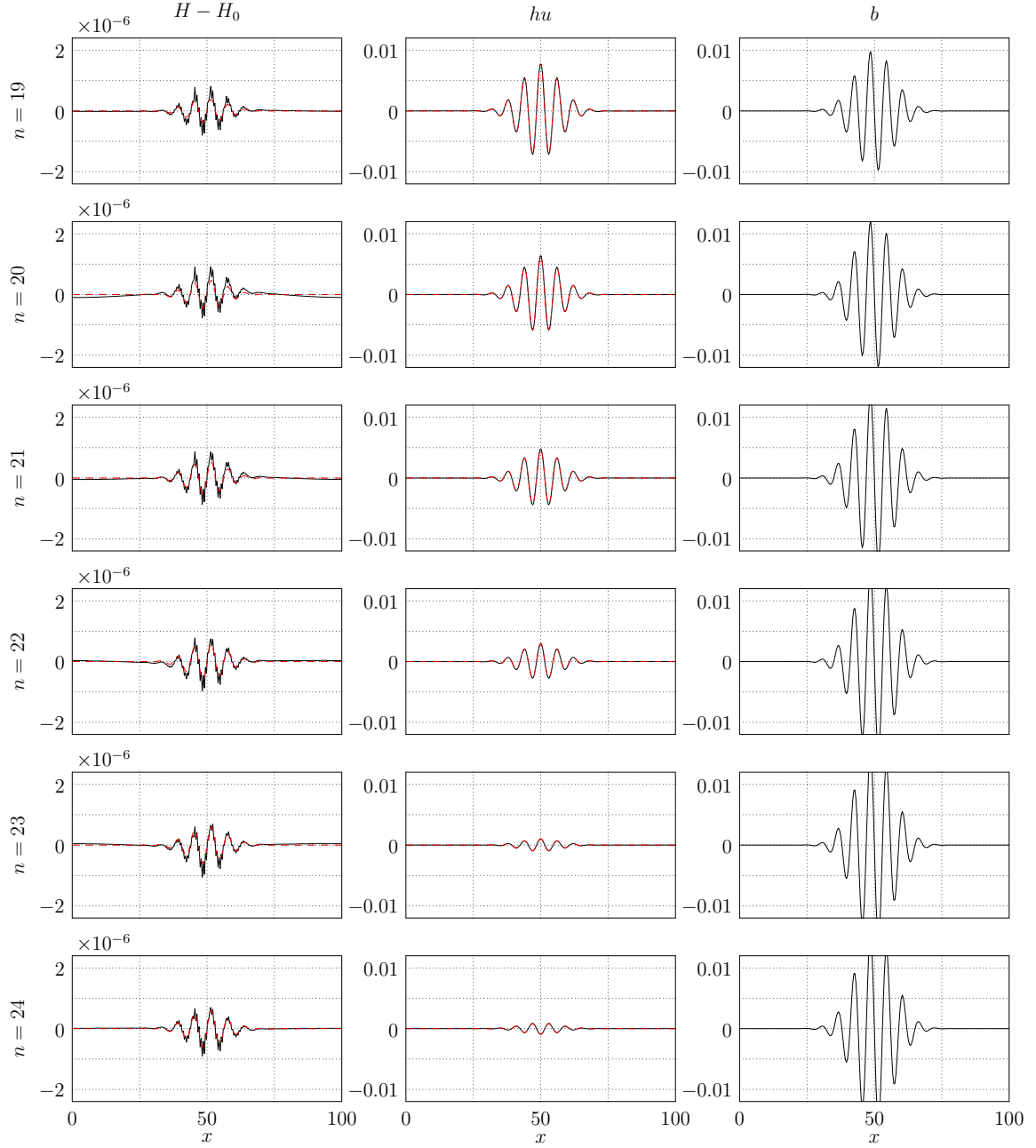
Figure 5.30: Same as Figure 5.29 with numerical solution after time steps 19 to 24.
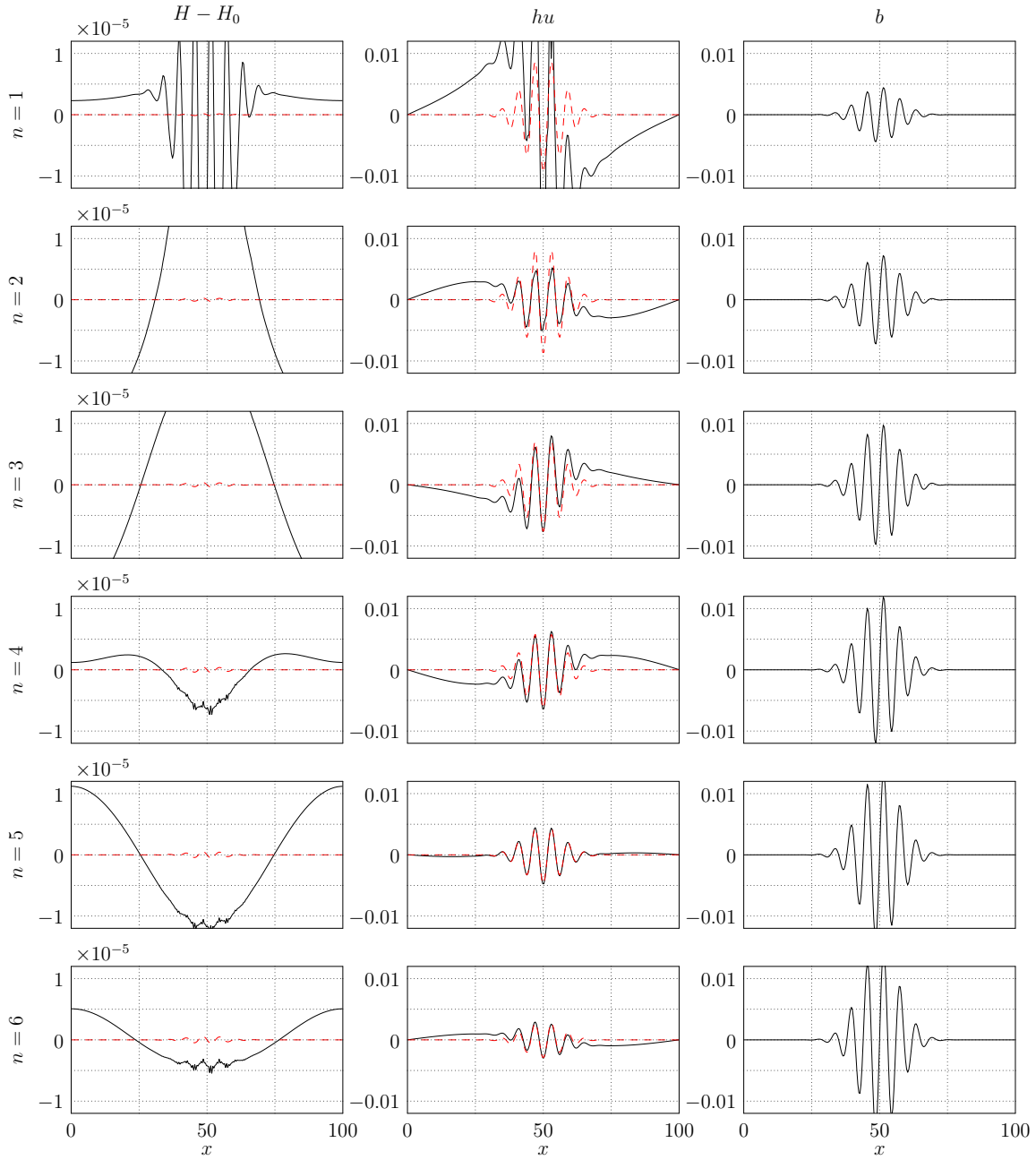
Figure 5.31: Completely unbalanced initial data. Numerical solution of the "multiscale test case" after the first six time steps using the scheme with BDF(2) in both corrections (black line) on a grid with 256 cells, Fr $= 0.01$. Asymptotic solution (from linear case) is plotted as dashed lines.
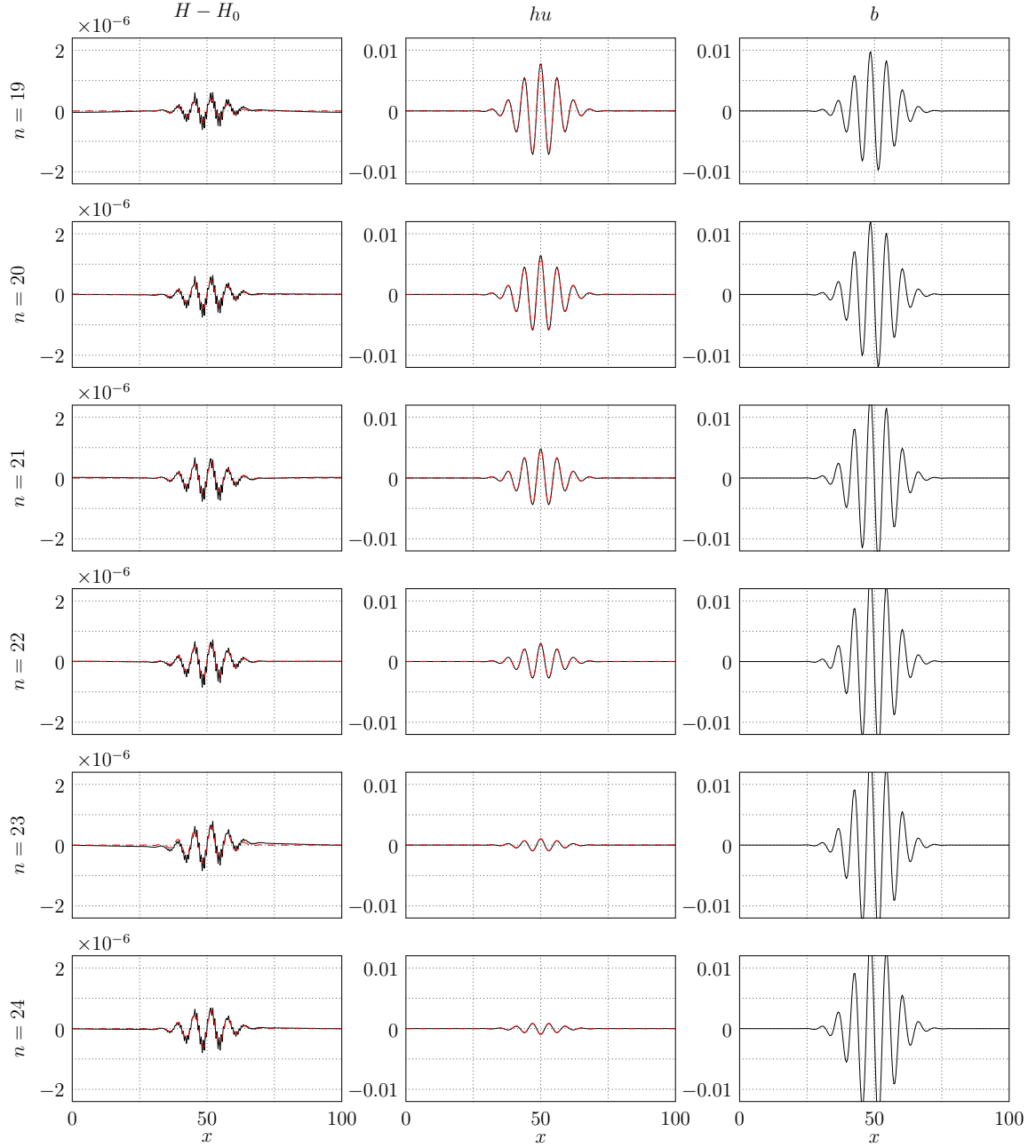
Figure 5.32: Same as Figure 5.31 with numerical solution after time steps 19 to 24.
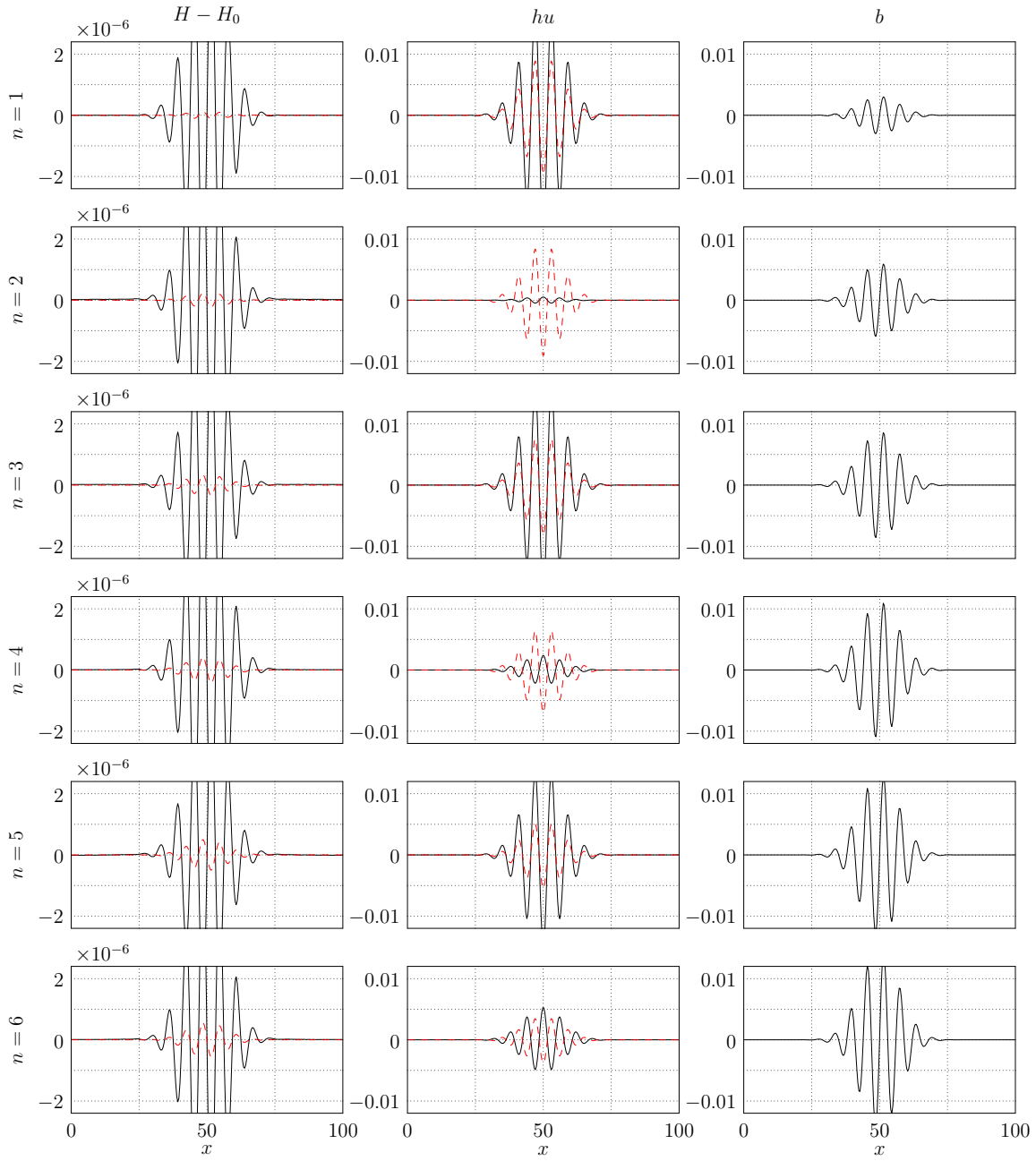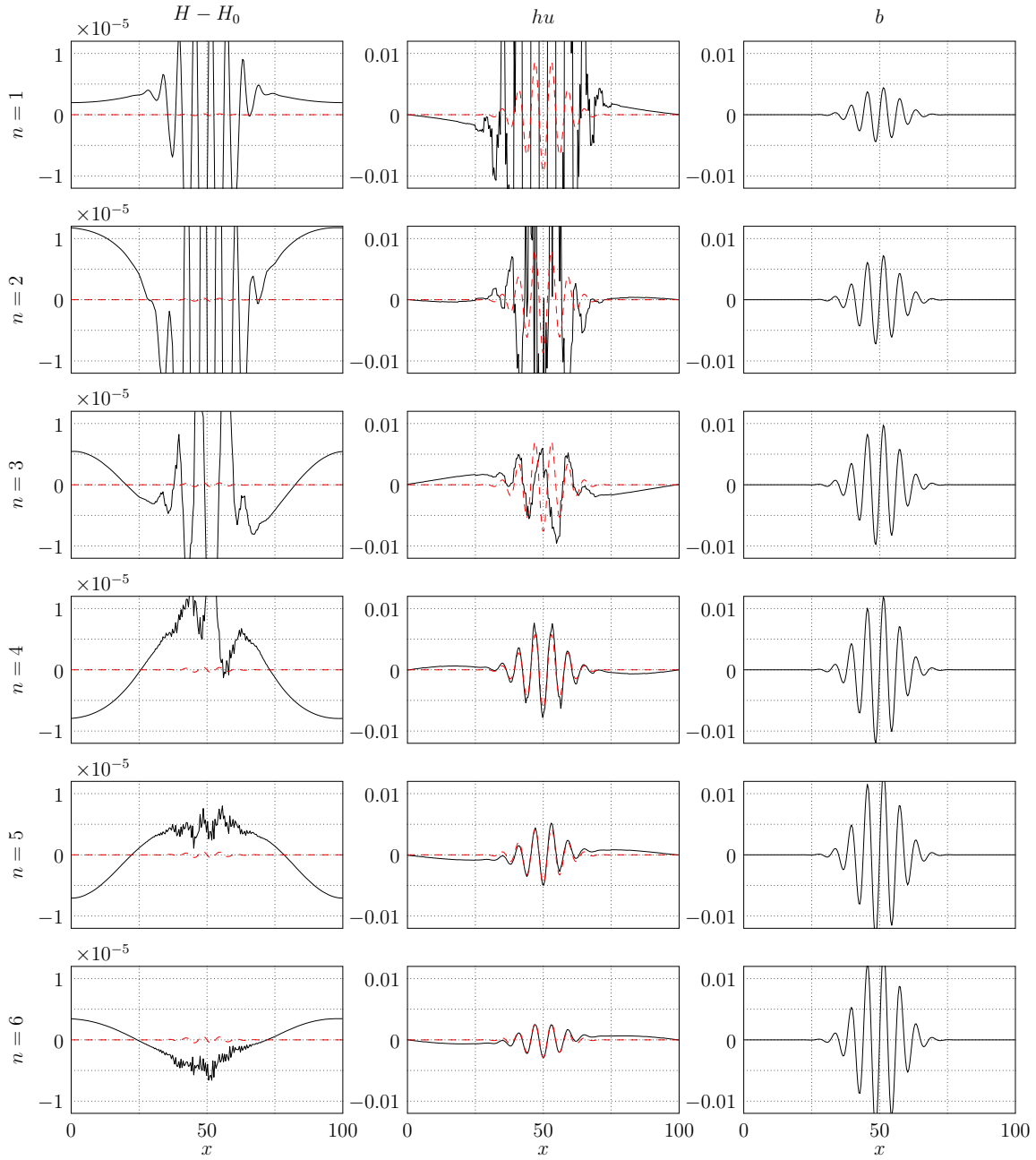
Figure 5.33: Completely unbalanced initial data. Numerical solution of the "multiscale test case" after the first six time steps using the multiscale scheme (black line) on a grid with 256 cells, Fr $= 0.01$. Asymptotic solution (from linear case) is plotted as dashed lines.
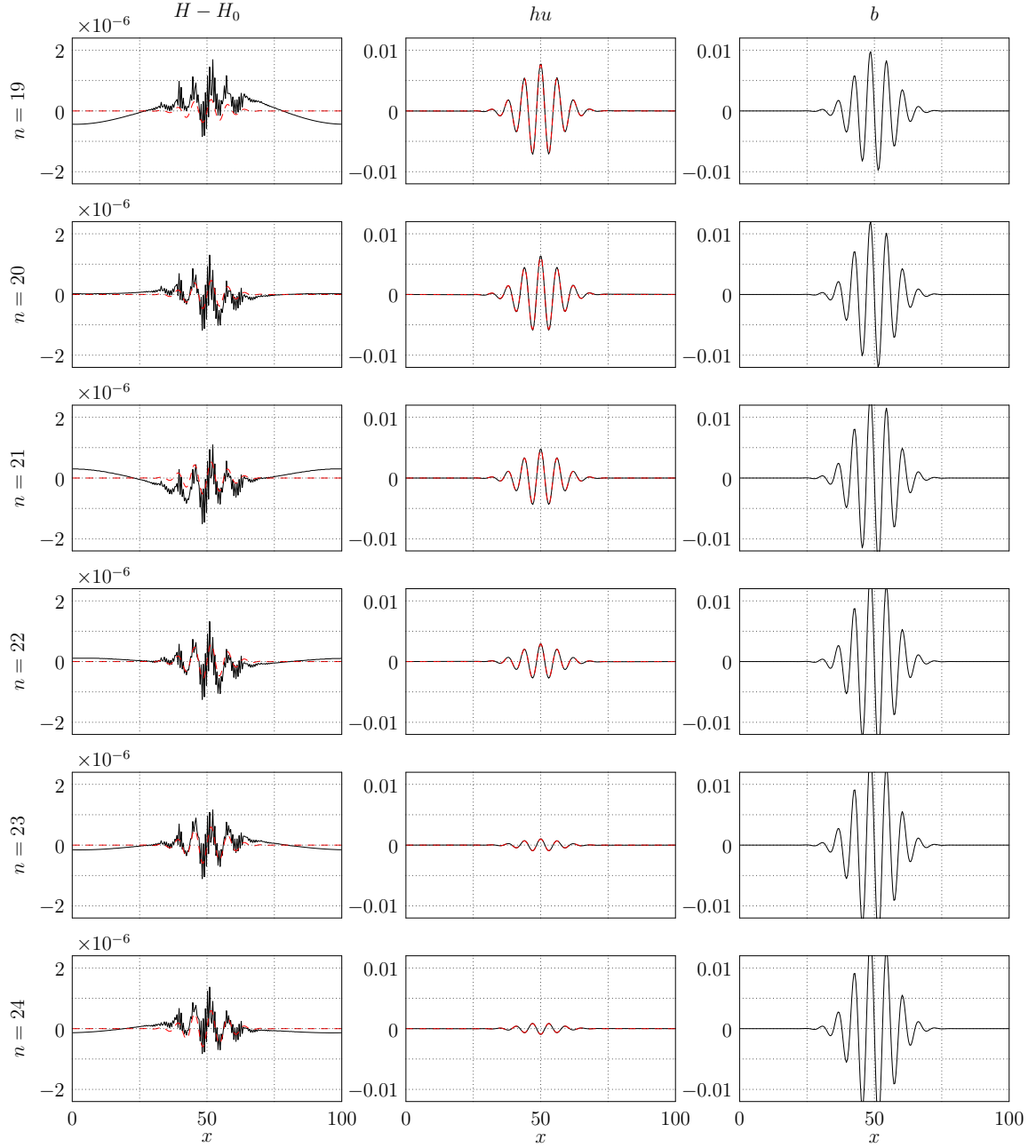
Figure 5.34: Same as Figure 5.33 with numerical solution after time steps 19 to 24.

6 Discussion

In this thesis, a new multiscale semi-implicit method for the numerical simulation of low Froude number shallow water flows was introduced. It is motivated by significant shortcomings of classical semi-implicit large time step integration schemes applied in current atmospheric codes. A principal feature of the new method is the diverse treatment of long and short wave data in accordance with the asymptotic regime of fast gravity waves traveling over short-range topography. This is achieved through a multilevel approach borrowing ideas from multigrid schemes for elliptic equations. The scheme is second-order accurate and admits time steps depending essentially on the flow velocity.

The following discussion provides a summary of the main findings and relates them to the objectives of this work. Furthermore, the numerical results obtained with the different schemes derived in the previous chapters are evaluated. The chapter concludes with an outlook for possible future research paths on this topic.

6.1 Discussion of the results

The multiscale scheme is based on two other methods, which were also derived in the course of this thesis. The first ingredient is a multilevel scheme for the solution of linearized shallow water flows, which was constructed in Chapter 3. It includes all the multiscale information of the discretization. The second building block is a projection method for the zero Froude number shallow water equations, presented in Chapter 4. This method provides a conservative discretization with a stable splitting between the slow and fast components of the flow that is consistent with the zero Froude number limit. These schemes are discussed separately, before the results of the semi-implicit method derived in Chapter 5 are reviewed.

Multilevel method for linearized shallow water flows

The numerical results for the multilevel method for linearized shallow water flows presented in Section 3.3 show that the scheme effectively eliminates freely propagating compressible short-wave components, which cannot be accurately represented at long time steps. At the same time, the dispersion and the amplitude errors for long-wave modes are minimized. In the presence of a source term that varies slowly in time but has rapid spatial variations, solutions relax to an asymptotic balanced state. This is consistent with the asymptotic regime of linearized long-wave shallow water waves traveling over topography as discussed in Section 2.2. In order to achieve these properties, a quasi-spectral decomposition of the data was performed using multigrid techniques, and a scale-dependent blending of two base schemes was applied.

Two classical implicit time step methods and a new super-implicit discretization were analyzed for the selection of the base schemes. The discrete-dispersion relations gave insight into the dispersion and amplitude errors produced depending on the wave number and the CFL number. Furthermore, the behavior of the schemes in presence of a source term was investigated. This analysis enabled us to effectively control the characteristics of the new method concerning the accurate transport of long-wave modes and dissipation of under-resolved short-wave gravity waves.

For the quasi-spectral decomposition of the height, classical multigrid prolongation and restriction operators were applied. This resulted in a good representation of the data associated with the different scales. On the other hand, the underlying staggered grid approach requires different transfer operators for the velocity than for the height, which complicates the implementation of the scheme. This might become even worse in case of more than one space dimension, where each component of the velocity probably needs to be treated differently in order to obtain an optimal match with the corresponding partial derivative of the height.

As pointed out, there is some arbitrariness in the determination of the weights which control the blending of the two base schemes depending on the scale. The presented approach using the implicit trapezoidal rule for grid levels with grid-CFL numbers smaller or equal to 1.0 and nudging the discretization towards a backward differentiation formula for the finer grid levels worked quite well in the numerical tests. Also, this is consistent with the asymptotic analysis in Section 2.2: At a grid level with grid-CFL number 1.0 the data should only include the long-wave component

$h^{(1)}$ of the height, which has only variation on the space scale $\xi = \mathsf{Fr}\,x$, on which the gravity wave speed is of order one. This component should be advected with minimized dispersion and no damping. On the other hand, in the extreme case where the CFL number on the fine grid is smaller or equal to 1.0, only the trapezoidal rule would be used by the multilevel scheme. Thus, high-wave-number modes that produce large dispersion errors would not be damped by the scheme. This shows that the choice of the blending weight might not be optimal yet.

Another open problem is the efficient numerical solution of the resulting Helmholtz equation in the multilevel method. As discussed in Section 3.2.4, the matrix representation of the discrete operator shows an unusual sparsity pattern that suggests that special numerical methods should be developed for the efficient solution of the system. Since the scale splitting already borrows ideas from multigrid methods, a promising approach would also go in this direction.

The overall idea of the scheme is quite general and can likely be applied also in other contexts in which only partially resolved processes require a scale-selective numerical representation. However, the current approach strongly relies on the linearity of the problem and the numerical base schemes employed.

Projection method for the lake equations

Since the general framework of the projection method for zero Froude number flows was already established in previous works (Vater, 2005; Vater and Klein, 2009), the focus was now set on the introduction of non-trivial time-dependent bottom topography. By using the pressureless equations with a source term as given in equation (4.2), no well-balancing due to the influence of bottom topography had to be considered in the predictor step. This greatly simplified the implementation of this part of the scheme. On the other hand, a degenerated hyperbolic system had to be solved. However, as it was pointed out in Section 2.5, the numerical solution of this system has been already studied in the literature (see, e.g., Bouchut *et al.*, 2003; LeVeque, 2004). The presented approach using second-order reconstruction techniques and the solution of the exact Riemann problem in the flux computation, resulted in stable and accurate results in case of the projection method.

The numerical results obtained with this method confirm its second-order accuracy and show a reasonable behavior in case of non-trivial bottom topography. The lake

at rest can be reproduced, even in the presence of small perturbations. In case of a vortex which is advected over topography principal features like the correct evolution of the maximum vorticity are well approximated for sufficiently fine grids.

Semi-implicit discretization of the shallow water equations
The extension of the projection method to a semi-implicit method for shallow water flows with non-zero Froude number was relatively straightforward. A key ingredient was the same splitting of the height variable as in the zero Froude number case into a background state and a second-order perturbation in terms of the Froude number. This led to the inclusion of the local time derivatives of this height perturbation. However, as we have seen in the numerical tests, this method can be triggered to produce unphysical oscillations, either in the weakly nonlinear regime where shocks occur, or in the presence of a source term which acts on the local divergence of the flow. This is due to a time discretization of the second correction equation, which is based on the implicit trapezoidal rule.

To overcome at least parts of these problems, two BDF(2)-based discretizations were introduced in Section 5.2. The design of these schemes has to be done with some care, because certain combinations of discretizations for the explicit and the implicit part can lead to unstable discretizations. Furthermore, the scheme with BDF(2)-based discretizations in both corrections resulted in a more restrictive CFL condition in the simulations. The choice of the particular discretization was guided by the results of the test cases in this work. However, it would be good if this choice could be grounded on more theoretical results analyzing the stability properties of semi-implicit BDF(2)-based discretizations.

All the above mentioned schemes were thoroughly tested concerning their convergence properties and the incorporation of bottom topography. The balancing properties of the schemes were tested by the simulation of the lake at rest. The results show that this state can be reproduced and the errors are essentially controlled by the accuracy of the linear solver used in the implementation of the schemes.

In the final step, a multiscale method was derived by combining the multilevel method for linearized shallow water flows with the semi-implicit schemes which are based on trapezoidal and BDF(2) discretizations in the second corrections, respectively. In this approach, only the second correction of the single-scale semi-implicit schemes

had to be modified by using the previously developed multilevel method for the linearized equations. The predictor step and the first correction were not affected by this. This method was implemented in one space dimension and with periodic boundary conditions in this work.

As in the linear case the multiscale scheme is able to properly advect long-wave gravity waves, and dispersion and amplitude errors are also minimized as much as the considered base schemes admit. However, some artifacts can be observed in the momentum field, which are probably related to the scale splitting within the operator. But these should be acceptable in practical applications. In the presence of bottom topography, which varies slowly in time, the balanced state is attained after a reasonable number of time steps.

In relation to Klein (1995) and Geratz (1997), where the equivalent asymptotic regime for the low Mach number Euler equations is considered, the presented scheme has a number of differences. While these authors explicitly distinguish three different pressure (height) components, in the present work a splitting into only two components is considered, and the acoustic pressure $\mathsf{M}p^{(1)}$ and local second-order pressure $\mathsf{M}^2p^{(2)}$ which takes care of the local balance of forces are not explicitly distinguished. In order to extract these different components, averages over areas of order $\Delta x/\mathsf{M}$ and $\Delta x/\mathsf{M}^2$, respectively, were computed. Furthermore, Klein and Geratz solved different equations for the acoustic and the second-order pressure. In the present discretization, the acoustic pressure is only implicitly included by the quasi-spectral decomposition of the data. The different roles of the long- and short-wave pressure components are accounted for by the scale-dependent blending of two different base discretizations.

Another difference is that in Klein and Geratz the acoustic and the second-order pressure are associated with the original fine grid. In contrast, by using the multigrid based splitting in the presented method, the acoustic pressure is implicitly associated with a coarse grid (where the grid CFL number is smaller or equal to one and the trapezoidal rule is used). This is more similar to Munz *et al.* (2003), for example, where the acoustic pressure equation was also solved on a coarse grid. However, these authors used the same averaging techniques as in Klein (1995), and they did not have any intermediate grid levels as in the present case.

In contrast to most other stabilization techniques used in codes for numerical weather prediction, the new scheme is still consistent of second-order accuracy. Other

techniques, like off-centering, divergence damping (Skamarock and Klemp, 1992) or subgrid scale (SGS) filters either reduce the order of the scheme to first order, or are not consistent with the underlying equations. Here, the truncation properties of certain "simple" schemes are deliberately employed to achieve positive effects in terms of subgrid-scale modeling. This is in the spirit of implicit Large Eddy Simulations (Grinstein *et al.*, 2007) for turbulence modeling.

As seen in the last two test cases, some problems concerning the predictor step of the scheme arise. In the simulation of a vortex, which is advected over topography, nonphysical oscillations can occur. Since these are probably due to insufficient numerical diffusion of the scheme it might be necessary to introduce some diffusion artificially in the solution of the Riemann problem, as it was done for example in Hickel *et al.* (2006). Also, the two stage Runge-Kutta scheme used in the predictor step is not able to correctly transport high-wave-number gravity wave modes. In the balancing test case with time dependent bottom topography this resulted in unphysical high-wave-number small-amplitude deviations. Therefore, either another numerical treatment of the "source term" in the auxiliary system might be needed, or another auxiliary system might have to be used which treats the gravity waves differently.

6.2 Conclusion and future research prospects

With the development of the multiscale method the goal of accurately simulating the asymptotic regime of fast gravity waves traveling over short range topography has been achieved. In the zero Froude number limit or for domains which are small compared to the large space scale associated with the propagation of gravity waves, the scheme switches into a projection method for zero Froude number flows. In this respect an asymptotically adaptive numerical method in the sense of Klein *et al.* (2001) was constructed.

Some open questions remain, however. For the selection of the base schemes, mostly classical discretizations were applied so far. There might be special discretizations more suitable for the particular purpose than the presented schemes. A first try in this direction was the construction of the super-implicit scheme. However, the results

from the test case with slow, short-wave source term suggest that this choice is not optimal. Furthermore, the weighting function μ_ν for the blending of the base schemes was chosen rather ad-hoc thus far, and more judicious choices aiming at optimization of the dispersion and damping effects are conceivable.

Since the multiscale scheme was only implemented in one space dimension, one of the next steps would be to extend the method to two space dimensions. This should be relatively straightforward, since the base schemes have been already thoroughly tested in this work. The only open task is the development of appropriate extensions of the transfer operators between the grid levels. Another important issue is the efficient implementation of the described method. Especially when applied to 2D or 3D problems, fast algorithms such as multigrid methods must be available. The challenging task in the development of such methods is to utilize the scale decomposition simultaneously for the blending of time integrators and for the efficient solution or preconditioning of the linear equations resulting from the semi-implicit discretization.

As was mentioned in the introduction, it is still not clear how much any acoustic mode really matters in atmospheric modeling. Thus, the importance of the proposed scheme might be debatable with respect to acoustics. However, as it was shown in Klein *et al.* (2010), there is also a clear scale separation between gravity wave modes and advection, and similar problems arise in this context. The presented numerical scheme (applied to the compressible flow equations) should be also applicable for this issue.

Finally, it should be mentioned that the numerical solution of the zero and low Froude number shallow water equations with non-trivial bottom topography is already valuable by itself for a number of research applications, such as the simulation of wave-driven currents near beaches (Bühler and Jacobson, 2001). To be able to address a wide rage of applications it is certainly desirable to include a discretization of the Coriolis term (see Audusse *et al.*, 2011, for a first step into this direction) and the capability of simulating beaches, i.e. vanishing water heights.

Appendix

A.1 Discrete-dispersion relations and amplification factors

As discussed in Section 3.1.4, the discrete-dispersion relation and amplification factor of a numerical scheme is obtained by substituting a traveling wave solution into the finite-difference formula. The staggered grid employed in the presented discretization leads to the ansatz

$$
\begin{aligned}
h_{j+1/2}^n &= \hat{h} \exp(-\mathrm{i}\omega n\Delta t) \exp(\mathrm{i}k(j + \tfrac{1}{2})\Delta x) \\
u_j^n &= \hat{u} \exp(-\mathrm{i}\omega n\Delta t) \exp(\mathrm{i}kj\Delta x)
\end{aligned}
\tag{A.1}
$$

for the discrete height and velocity variables at time t^n and position $x_{j+1/2}$ for the height and x_j for the velocity. This leads to a relationship between the spatial wave number k and the frequency ω, where $\omega = \omega_\mathrm{r} + \mathrm{i}\omega_\mathrm{i}$ is a complex number.

Implicit trapezoidal rule

Therefore, for the analysis of the implicit trapezoidal rule, (A.1) is inserted into

$$
\begin{aligned}
\frac{1}{\Delta t}\left(u_j^{n+1} - u_j^n\right) &= -\frac{1}{2\mathsf{Fr}^2\Delta x}\left(h_{j+1/2}^{n+1} - h_{j-1/2}^{n+1} + h_{j+1/2}^n - h_{j-1/2}^n\right) , \\
\frac{1}{\Delta t}\left(h_{j+1/2}^{n+1} - h_{j+1/2}^n\right) &= -\frac{H_0}{2\Delta x}\left(u_{j+1}^{n+1} - u_{j-1}^{n+1} + u_{j+1}^n - u_{j-1}^n\right) .
\end{aligned}
$$

From this one obtains after division by the common exponential factors

$$
\begin{aligned}
[1 - \exp(\mathrm{i}\omega\Delta t)]\,\hat{u} + \frac{\Delta t}{\mathsf{Fr}^2\Delta x}\,[1 + \exp(\mathrm{i}\omega\Delta t)] \cdot \left[\exp(\mathrm{i}\tfrac{k}{2}\Delta x) - \exp(-\mathrm{i}\tfrac{k}{2}\Delta x)\right]\hat{h} &= 0 , \\
[1 - \exp(\mathrm{i}\omega\Delta t)]\,\hat{h} + \frac{H_0\Delta t}{2\Delta x}\,[1 + \exp(\mathrm{i}\omega\Delta t)] \cdot \left[\exp(\mathrm{i}\tfrac{k}{2}\Delta x) - \exp(-\mathrm{i}\tfrac{k}{2}\Delta x)\right]\hat{u} &= 0 .
\end{aligned}
$$

161

Since $2\mathrm{i}\sin(\frac{k}{2}\Delta x) = \exp(\mathrm{i}\frac{k}{2}\Delta x) - \exp(-\mathrm{i}\frac{k}{2}\Delta x)$ and by definition of $\xi \coloneqq \exp(\mathrm{i}\omega\Delta t)$ this leads to the linear system

$$\begin{pmatrix} 1 - \xi & \mathrm{i}\frac{H_0\Delta t}{\Delta x}\sin(\frac{k}{2}\Delta x) \\ \mathrm{i}\frac{\Delta t}{\mathsf{Fr}^2\Delta x}\sin(\frac{k}{2}\Delta x) & 1 - \xi \end{pmatrix} \begin{pmatrix} \hat{h} \\ \hat{u} \end{pmatrix} = 0 \;.$$

To determine the non-trivial solutions of this system, let us analyze for which ξ the determinant of the system matrix is 0. With $\alpha \coloneqq \frac{\sqrt{H_0}\Delta t}{\mathsf{Fr}\Delta x}\sin(\frac{k}{2}\Delta x)$ this means that ξ must satisfy

$$(1-\xi)^2 + \alpha^2(1+\xi)^2 = 0 \;.$$

The solution of this equation is

$$\xi_{1/2} = \frac{(1\pm\mathrm{i}\alpha)^2}{\alpha^2 + 1} \;,$$

which has an absolute value of one for all α. Due to the definition of ξ, ω must then be real (i.e., $\omega_\mathrm{i} = 0$), and the amplitude is kept constant in time for all wave numbers k. This means that the amplification factor is given by $A \coloneqq \exp(\omega_\mathrm{i}\Delta t) = 1$.

The phase error is obtained by solving $\xi = \exp(\mathrm{i}\omega_\mathrm{r}\Delta t)$ for ω_r, which leads to

$$\omega_\mathrm{r} = \pm\frac{2}{\Delta t}\arctan\left(\frac{\sqrt{H_0}\Delta t}{\mathsf{Fr}\Delta x}\cdot\sin\left(\frac{k\Delta x}{2}\right)\right) \;.$$

BDF(2) scheme

For the BDF(2) scheme the ansatz (A.1) has to be inserted into

$$\frac{1}{\Delta t}\left(u_j^{n+1} - \tfrac{4}{3}u_j^n + \tfrac{1}{3}u_j^{n-1}\right) = -\frac{2}{3\mathsf{Fr}^2\Delta x}\left(h_{j+1/2}^{n+1} - h_{j-1/2}^{n+1}\right) \;,$$

$$\frac{1}{\Delta t}\left(h_{j+1/2}^{n+1} - \tfrac{4}{3}h_{j+1/2}^n + \tfrac{1}{3}h_{j+1/2}^{n-1}\right) = -\frac{2H_0}{3\Delta x}\left(u_{j+1}^{n+1} - u_{j-1}^{n+1}\right) \;,$$

to obtain after division by the common exponential factors

$$\left[1 - \tfrac{4}{3}\exp(\mathrm{i}\omega\Delta t) + \tfrac{1}{3}\exp(2\mathrm{i}\omega\Delta t)\right]\hat{u} + \frac{2\Delta t}{3\mathsf{Fr}^2\Delta x}\left[\exp(\mathrm{i}\tfrac{k}{2}\Delta x) - \exp(-\mathrm{i}\tfrac{k}{2}\Delta x)\right]\hat{h} = 0 \;,$$

$$\left[1 - \tfrac{4}{3}\exp(\mathrm{i}\omega\Delta t) + \tfrac{1}{3}\exp(2\mathrm{i}\omega\Delta t)\right]\hat{h} + \frac{2H_0\Delta t}{3\Delta x}\left[\exp(\mathrm{i}\tfrac{k}{2}\Delta x) - \exp(-\mathrm{i}\tfrac{k}{2}\Delta x)\right]\hat{u} = 0 \;.$$

By the same definitions as for the trapezoidal rule this leads to the linear system

$$
\begin{pmatrix} 1 - \frac{4}{3}\xi + \frac{1}{3}\xi^2 & \mathrm{i}\frac{4H_0\Delta t}{3\Delta x}\sin(\frac{k}{2}\Delta x) \\ \mathrm{i}\frac{4\Delta t}{3\mathsf{Fr}^2\Delta x}\sin(\frac{k}{2}\Delta x) & 1 - \frac{4}{3}\xi + \frac{1}{3}\xi^2 \end{pmatrix} \begin{pmatrix} \hat{h} \\ \hat{u} \end{pmatrix} = 0 \;.
$$

To obtain non-trivial solutions of this system, let $\alpha := \frac{\sqrt{H_0}\Delta t}{\mathsf{Fr}\Delta x}\sin(\frac{k}{2}\Delta x)$. Then, ξ must satisfy

$$
\left(1 - \tfrac{4}{3}\xi + \tfrac{1}{3}\xi^2\right)^2 = -\left(\tfrac{4}{3}\alpha\right)^2 \;,
$$

which is the case, if ξ further satisfies

$$
1 - \tfrac{4}{3}\xi + \tfrac{1}{3}\xi^2 \pm \tfrac{4}{3}\mathrm{i}\alpha = 0 \;.
$$

The solution of this complex quadratic equation is given by

$$
\begin{aligned}
\xi_{1/2} &= 2 \pm \sqrt{1 \pm 4\mathrm{i}\alpha} \\
&= \left[2 \pm \sqrt[4]{1 + 16\alpha^2}\cos\left(\tfrac{1}{2}\arctan(4\alpha)\right)\right] \pm \mathrm{i}\left[\sqrt[4]{1 + 16\alpha^2}\sin\left(\tfrac{1}{2}\arctan(4\alpha)\right)\right] \;.
\end{aligned}
$$

By taking the absolute value, this yields the amplification factor, which is

$$
A = \frac{1}{|\xi|} = \left(4 \pm 4\sqrt[4]{1 + 16\alpha^2}\cos\left(\tfrac{1}{2}\arctan(4\alpha)\right) + \sqrt[4]{1 + 16\alpha^2}\right)^{-1/2}
$$

with $\alpha := \frac{\sqrt{H_0}\Delta t}{\mathsf{Fr}\Delta x}\sin(\frac{k}{2}\Delta x)$, as before. The discrete-dispersion relation is given by

$$
\omega_{\mathrm{r}} = \pm\frac{1}{\Delta t}\arctan\left(\frac{\sqrt[4]{1 + 16\alpha^2}\sin\left(\tfrac{1}{2}\arctan(4\alpha)\right)}{2 \pm \sqrt[4]{1 + 16\alpha^2}\cos\left(\tfrac{1}{2}\arctan(4\alpha)\right)}\right) \;.
$$

To determine which is the physical and which is the computational mode, the amplification factor can be further analyzed. For this, let us consider the limit $\Delta t, \Delta x \to 0$ for $\Delta t/\Delta x = \mathrm{const}$. For a convergent method the amplification factor should tend to one in this limit. Since $\alpha \to 0$ for $\Delta x \to 0$, this leads to

$$
A \to (4 \pm 4 + 1)^{-1/2} = \{\tfrac{1}{3}, 1\} \quad \text{for } \alpha \to 0 \;,
$$

and the $(-)$-branch represents the physical mode.

A.2 Convergence results for the zero Froude number projection method

A.2.1 "Flow over a hill" test case

Table A.1: Experimental errors and convergence rates in $(h\boldsymbol{u})$ for the "flow over a hill" test case for different choices of v_{rel} at time $t_{\text{end}} = 10$.

test case	norm	32×32	rate γ	64×64	rate γ	128×128	rate γ	256×256
$v_{\text{rel}} = -1.0$	L^2	6.9452e-03	2.376	1.3378e-03	2.654	2.1251e-04	2.526	3.6902e-05
	L^∞	1.7706e-02	2.165	3.9490e-03	2.608	6.4779e-04	2.776	9.4607e-05
$v_{\text{rel}} = -0.5$	L^2	2.7995e-02	1.088	1.3168e-02	1.710	4.0242e-03	1.953	1.0397e-03
	L^∞	6.4006e-02	0.923	3.3761e-02	1.601	1.1129e-02	1.945	2.8902e-03
$v_{\text{rel}} = 0.0$	L^2	3.1491e-02	1.033	1.5392e-02	1.688	4.7755e-03	1.989	1.2031e-03
	L^∞	7.2840e-02	0.995	3.6549e-02	1.524	1.2712e-02	1.925	3.3485e-03
$v_{\text{rel}} = +1.0$	L^2	4.7805e-02	1.232	2.0354e-02	1.922	5.3730e-03	2.622	8.7309e-04
	L^∞	1.0009e-01	1.039	4.8725e-02	1.729	1.4699e-02	2.384	2.8166e-03

A.3 Convergence results for the low Froude number semi-implicit scheme

A.3.1 "Simple wave" test case

Table A.2: Errors and convergence rates in h for the different variants of the semi-implicit method. One-dimensional simple wave test case.

method	norm	256	rate γ	512	rate γ	1024	rate γ	2048
standard scheme	L^2	3.2801e-03	1.846	9.1251e-04	1.955	2.3530e-04	1.991	5.9190e-05
	L^∞	1.0686e-02	1.705	3.2770e-03	1.898	8.7942e-04	1.977	2.2342e-04
BDF(2) only in 2nd correction	L^2	4.7937e-03	1.763	1.4127e-03	1.912	3.7548e-04	1.975	9.5495e-05
	L^∞	1.4599e-02	1.587	4.8593e-03	1.822	1.3743e-03	1.947	3.5642e-04
BDF(2) in both corrections	L^2	6.6538e-03	1.589	2.2110e-03	1.803	6.3344e-04	1.925	1.6685e-04
	L^∞	1.8923e-02	1.413	7.1074e-03	1.681	2.2162e-03	1.867	6.0747e-04
multiscale method	L^2	3.2929e-03	1.849	9.1434e-04	1.957	2.3547e-04	1.992	5.9202e-05
	L^∞	1.0743e-02	1.712	3.2800e-03	1.896	8.8151e-04	1.978	2.2381e-04

Table A.3: Errors and convergence rates in (hu) for the different variants of the semi-implicit method. One-dimensional simple wave test case.

method	norm	256	rate γ	512	rate γ	1024	rate γ	2048
standard scheme	L^2	3.2422e-02	1.864	8.9047e-03	1.961	2.2875e-03	1.991	5.7556e-04
	L^∞	1.0527e-01	1.722	3.1899e-02	1.904	8.5226e-03	1.977	2.1654e-03
BDF(2) only in 2nd correction	L^2	4.7676e-02	1.740	1.4277e-02	1.910	3.8002e-03	1.976	9.6614e-04
	L^∞	1.4534e-01	1.573	4.8843e-02	1.826	1.3778e-02	1.952	3.5620e-03
BDF(2) in both corrections	L^2	6.8490e-02	1.620	2.2283e-02	1.820	6.3107e-03	1.934	1.6511e-03
	L^∞	1.9402e-01	1.436	7.1715e-02	1.700	2.2071e-02	1.879	5.9999e-03
multiscale method	L^2	3.2567e-02	1.867	8.9255e-03	1.963	2.2893e-03	1.992	5.7568e-04
	L^∞	1.0883e-01	1.758	3.2177e-02	1.901	8.6172e-03	1.984	2.1783e-03

A.3.2 "2D convergence" test case

Table A.4: Errors and convergence rates in h for the different variants of the semi-implicit method. 2D convergence test case.

method	norm	64x64	rate γ	128x128	rate γ	256x256	rate γ	512x512
standard scheme	L^2	9.9064e-05	1.909	2.6371e-05	2.007	6.5628e-06	1.994	1.6480e-06
	L^∞	3.8147e-04	1.970	9.7355e-05	2.079	2.3045e-05	2.043	5.5916e-06
BDF(2) only in 2nd correction	L^2	9.7523e-05	1.885	2.6411e-05	2.010	6.5567e-06	1.992	1.6480e-06
	L^∞	3.7566e-04	2.000	9.3911e-05	2.064	2.2466e-05	1.997	5.6298e-06
BDF(2) in both corrections	L^2	9.6838e-05	1.868	2.6538e-05	2.014	6.5695e-06	1.999	1.6435e-06
	L^∞	3.7470e-04	1.979	9.5062e-05	2.061	2.2788e-05	2.015	5.6387e-06

Table A.5: Errors and convergence rates in $(h\boldsymbol{u})$ for the different variants of the semi-implicit method. 2D convergence test case.

method	norm	64x64	rate γ	128x128	rate γ	256x256	rate γ	512x512
standard scheme	L^2	2.6095e-02	1.689	8.0912e-03	1.909	2.1545e-03	1.976	5.4748e-04
	L^∞	8.6873e-02	1.553	2.9596e-02	2.065	7.0742e-03	2.021	1.7436e-03
BDF(2) only in 2nd correction	L^2	2.6093e-02	1.689	8.0912e-03	1.909	2.1545e-03	1.976	5.4748e-04
	L^∞	8.6864e-02	1.553	2.9607e-02	2.065	7.0769e-03	2.022	1.7429e-03
BDF(2) in both corrections	L^2	2.6095e-02	1.689	8.0912e-03	1.909	2.1545e-03	1.976	5.4751e-04
	L^∞	8.7533e-02	1.556	2.9766e-02	2.066	7.1081e-03	2.031	1.7397e-03

A.3.3 "Flow over a hill" test case

Table A.6: Experimental errors and convergence rates in h for the "flow over a hill" test case for $v_{\mathrm{rel}} = -1.0$ at time $t_{\mathrm{end}} = 0.1$.

method	norm	256	rate	512	rate γ	1024	rate γ	2048
standard scheme	L^2	9.0784e-07	2.335	1.7999e-07	2.244	3.7995e-08	2.072	9.0336e-09
	L^∞	2.3417e-06	2.260	4.8884e-07	2.232	1.0404e-07	2.128	2.3795e-08
BDF(2) only in 2nd correction	L^2	1.5132e-06	2.415	2.8377e-07	2.419	5.3060e-08	2.206	1.1504e-08
	L^∞	3.7677e-06	2.383	7.2230e-07	2.338	1.4290e-07	2.209	3.0913e-08
BDF(2) in both corrections	L^2	1.4698e-06	2.179	3.2453e-07	2.257	6.7882e-08	2.143	1.5366e-08
	L^∞	4.1586e-06	2.122	9.5532e-07	2.283	1.9629e-07	2.283	4.0324e-08
multilevel method	L^2	9.2448e-07	2.323	1.8480e-07	2.229	3.9424e-08	2.066	9.4143e-09
	L^∞	2.4409e-06	2.164	5.4453e-07	2.208	1.1789e-07	2.111	2.7282e-08

Table A.7: Experimental errors and convergence rates in (hu) for the "flow over a hill" test case for $v_{\mathrm{rel}} = -1.0$ at time $t_{\mathrm{end}} = 0.1$.

method	norm	256	rate	512	rate γ	1024	rate γ	2048
standard scheme	L^2	3.3229e-06	1.615	1.0849e-06	1.791	3.1341e-07	1.968	8.0137e-08
	L^∞	6.6023e-06	1.177	2.9200e-06	1.839	8.1640e-07	2.007	2.0307e-07
BDF(2) only in 2nd correction	L^2	6.9586e-06	3.062	8.3346e-07	1.560	2.8272e-07	1.867	7.7505e-08
	L^∞	1.8673e-05	3.291	1.9081e-06	1.335	7.5640e-07	1.933	1.9804e-07
BDF(2) in both corrections	L^2	1.9227e-05	1.858	5.3057e-06	1.935	1.3872e-06	1.992	3.4876e-07
	L^∞	4.7157e-05	1.847	1.3112e-05	1.847	3.6440e-06	1.964	9.3387e-07
multilevel method	L^2	3.3666e-06	1.634	1.0850e-06	1.791	3.1348e-07	1.968	8.0143e-08
	L^∞	7.8736e-06	1.394	2.9953e-06	1.866	8.2160e-07	2.011	2.0389e-07

Table A.8: Experimental errors and convergence rates in h for the "flow over a hill" test case for $v_{\text{rel}} = -0.5$ at time $t_{\text{end}} = 0.1$.

method	norm	256	rate	512	rate γ	1024	rate γ	2048
standard scheme	L^2	7.1980e-07	2.115	1.6611e-07	2.050	4.0101e-08	2.013	9.9347e-09
	L^∞	1.7959e-06	2.090	4.2196e-07	2.067	1.0067e-07	2.035	2.4559e-08
BDF(2) only in 2nd correction	L^2	8.6551e-07	2.196	1.8895e-07	2.125	4.3304e-08	2.042	1.0514e-08
	L^∞	2.2123e-06	2.190	4.8496e-07	2.128	1.1092e-07	2.071	2.6396e-08
BDF(2) in both corrections	L^2	9.8900e-07	2.133	2.2544e-07	2.190	4.9411e-08	2.086	1.1638e-08
	L^∞	2.4749e-06	2.097	5.7858e-07	2.178	1.2786e-07	2.147	2.8861e-08
multiscale method	L^2	7.2290e-07	2.117	1.6667e-07	2.051	4.0214e-08	2.014	9.9544e-09
	L^∞	1.9098e-06	2.074	4.5362e-07	2.072	1.0789e-07	2.051	2.6041e-08

Table A.9: Experimental errors and convergence rates in (hu) for the "flow over a hill" test case for $v_{\text{rel}} = -0.5$ at time $t_{\text{end}} = 0.1$.

method	norm	256	rate	512	rate γ	1024	rate γ	2048
standard scheme	L^2	2.2528e-06	2.347	4.4291e-07	2.001	1.1064e-07	2.009	2.7494e-08
	L^∞	4.9717e-06	2.129	1.1367e-06	2.140	2.5793e-07	1.979	6.5409e-08
BDF(2) only in 2nd correction	L^2	3.9025e-06	3.047	4.7201e-07	2.125	1.0820e-07	1.991	2.7221e-08
	L^∞	9.4629e-06	2.888	1.2780e-06	2.273	2.6441e-07	2.071	6.2910e-08
BDF(2) in both corrections	L^2	1.0867e-05	1.908	2.8956e-06	1.955	7.4679e-07	1.995	1.8740e-07
	L^∞	2.7461e-05	1.963	7.0419e-06	1.950	1.8221e-06	1.980	4.6192e-07
multiscale method	L^2	2.2745e-06	2.358	4.4379e-07	2.003	1.1070e-07	2.009	2.7498e-08
	L^∞	5.7453e-06	2.219	1.2339e-06	2.210	2.6660e-07	2.012	6.6104e-08

Table A.10: Experimental errors and convergence rates in h for the "flow over a hill" test case for $v_{\text{rel}} = 0.0$ at time $t_{\text{end}} = 0.1$.

method	norm	256	rate	512	rate γ	1024	rate γ	2048
standard scheme	L^2	7.5100e-07	1.921	1.9835e-07	2.000	4.9572e-08	1.999	1.2404e-08
	L^∞	2.1401e-06	1.922	5.6480e-07	2.148	1.2747e-07	2.051	3.0760e-08
BDF(2) only in 2nd correction	L^2	6.9085e-07	1.797	1.9880e-07	2.002	4.9645e-08	2.001	1.2405e-08
	L^∞	1.9452e-06	1.725	5.8853e-07	2.144	1.3319e-07	2.088	3.1319e-08
BDF(2) in both corrections	L^2	6.1168e-07	1.689	1.8973e-07	1.942	4.9377e-08	1.995	1.2390e-08
	L^∞	1.6894e-06	1.623	5.4862e-07	2.011	1.3607e-07	2.109	3.1550e-08
multiscale method	L^2	7.5006e-07	1.919	1.9836e-07	2.001	4.9572e-08	1.999	1.2404e-08
	L^∞	2.1370e-06	1.919	5.6501e-07	2.148	1.2747e-07	2.051	3.0761e-08

Table A.11: Experimental errors and convergence rates in (hu) for the "flow over a hill" test case for $v_{\mathrm{rel}} = 0.0$ at time $t_{\mathrm{end}} = 0.1$.

method	norm	256	rate	512	rate γ	1024	rate γ	2048
standard scheme	L^2	3.0622e-06	1.399	1.1609e-06	1.899	3.1119e-07	1.987	7.8498e-08
	L^∞	8.2325e-06	1.688	2.5550e-06	1.882	6.9321e-07	1.985	1.7505e-07
BDF(2) only in 2nd correction	L^2	2.4831e-06	1.267	1.0316e-06	1.756	3.0537e-07	1.964	7.8272e-08
	L^∞	7.2760e-06	1.616	2.3739e-06	1.812	6.7587e-07	1.954	1.7447e-07
BDF(2) in both corrections	L^2	2.6358e-06	1.600	8.6930e-07	1.562	2.9437e-07	1.919	7.7822e-08
	L^∞	7.6952e-06	1.812	2.1915e-06	1.769	6.4306e-07	1.893	1.7320e-07
multiscale method	L^2	3.0403e-06	1.390	1.1602e-06	1.899	3.1118e-07	1.987	7.8498e-08
	L^∞	8.1731e-06	1.678	2.5540e-06	1.881	6.9319e-07	1.985	1.7505e-07

Table A.12: Experimental errors and convergence rates in h for the "flow over a hill" test case for $v_{\mathrm{rel}} = 1.0$ at time $t_{\mathrm{end}} = 0.1$.

method	norm	256	rate	512	rate γ	1024	rate γ	2048
standard scheme	L^2	1.7290e-06	1.908	4.6084e-07	2.027	1.1307e-07	2.006	2.8153e-08
	L^∞	4.6636e-06	1.756	1.3807e-06	2.211	2.9822e-07	2.087	7.0174e-08
BDF(2) only in 2nd correction	L^2	1.8612e-06	1.846	5.1779e-07	2.062	1.2400e-07	2.031	3.0334e-08
	L^∞	4.5188e-06	1.556	1.5368e-06	2.147	3.4696e-07	2.146	7.8393e-08
BDF(2) in both corrections	L^2	9.8138e-07	1.231	4.1821e-07	1.704	1.2832e-07	1.933	3.3606e-08
	L^∞	2.5477e-06	0.974	1.2967e-06	1.780	3.7748e-07	2.103	8.7853e-08
multilevel method	L^2	1.7301e-06	1.907	4.6129e-07	2.028	1.1310e-07	2.006	2.8155e-08
	L^∞	4.7920e-06	1.773	1.4021e-06	2.218	3.0139e-07	2.093	7.0626e-08

Table A.13: Experimental errors and convergence rates in (hu) for the "flow over a hill" test case for $v_{\mathrm{rel}} = 1.0$ at time $t_{\mathrm{end}} = 0.1$.

method	norm	256	rate	512	rate γ	1024	rate γ	2048
standard scheme	L^2	9.3609e-06	1.672	2.9368e-06	1.900	7.8697e-07	1.990	1.9810e-07
	L^∞	2.7361e-05	1.870	7.4866e-06	1.950	1.9373e-06	2.033	4.7331e-07
BDF(2) only in 2nd correction	L^2	1.0527e-05	1.997	2.6375e-06	1.781	7.6768e-07	1.957	1.9767e-07
	L^∞	2.7558e-05	2.024	6.7776e-06	1.877	1.8454e-06	1.983	4.6698e-07
BDF(2) in both corrections	L^2	1.5410e-05	1.393	5.8695e-06	1.694	1.8140e-06	1.933	4.7511e-07
	L^∞	4.0458e-05	1.409	1.5240e-05	1.631	4.9189e-06	1.915	1.3047e-06
multilevel method	L^2	9.3512e-06	1.672	2.9349e-06	1.899	7.8699e-07	1.990	1.9810e-07
	L^∞	2.7653e-05	1.874	7.5438e-06	1.953	1.9484e-06	2.037	4.7468e-07

A.4 Zusammenfassung

In dieser Arbeit wird ein neues semi-implizites Mehrskalenverfahren zur Berechnung von Flachwasserströmungen bei kleinen Froudezahlen vorgestellt. Motiviert durch meteorologische Anwendungen zielt es darauf ab, Dispersions- und Amplitudenfehler bei der Berechnung von langwelligen Schwerewellen zu minimieren. Während die durch langsame Anregung aufgezwungene Dynamik in kurzwelligen Lösungskomponenten korrekt balanciert wird, unterdrückt das Verfahren sich frei ausbreitende kurzwellige Moden, die in der Zeit nicht aufgelöst sind. Dies wird durch einen Multilevelansatz erreicht, der sich Konzepten bedient, die von Mehrgitterverfahren zur Lösung von elliptischen Gleichungen bekannt sind. Das Verfahren ist von zweiter Ordnung und erlaubt Zeitschritte, die nur von der Strömungsgeschwindigkeit abhängen.

In einem ersten Schritt wird ein Multilevelverfahren zur Lösung der eindimensionalen linearisierten Flachwassergleichungen hergeleitet. Die Zerlegung der Daten nach verschiedenen Skalen ermöglicht eine skalenabhängige Verschneidung von Zeitintegratoren mit verschiedenen wesentlichen Merkmalen. Zur Auswahl der Integratoren werden die diskreten Dispersionsrelationen einiger klassischer Zweite-Ordnung-Verfahren berechnet und deren Verhalten im Falle von niederfrequenten räumlich stark variierenden Quelltermen diskutiert. Das resultierende Verfahren besteht im Wesentlichen aus der Lösung eines Helmholtzproblems auf dem ursprünglichen feinen Gitter, wobei der Differenzenoperator und die rechte Seite die Mehrskaleninformation der Diskretisierung beinhalten. Die Güte des Verfahrens wird in einem Testfall mit „mehrskaligen" Anfangsdaten analysiert. Ein weiterer Testfall gibt Aufschluss über das Verhalten in Anwesenheit eines sich langsam in der Zeit ändernden Quellterms, der räumlich stark variiert.

Das Verfahren zur Simulation der vollen nichtlinearen Flachwassergleichungen baut auf ein Projektionsverfahren für die Grenzgleichungen im Limes Froudezahl gegen Null auf. Dafür wird die in Vater und Klein (*Numer. Math.* 113, S. 123–161, 2009) beschriebene Methode auf Probleme mit zeitlich variierender Bodentopographie erweitert. Numerische Simulationen zeigen, dass das Verfahren gut-balanciert ist und einen ruhenden See mit nichttrivialer Bodentopographie berechnen kann. Die Ergebnisse zwei weiterer Testfälle bestätigen die korrekte Darstellung der zeitabhängigen Bodentopographie.

In einem letzten Schritt wird das semi-implizite Mehrskalenverfahren für die numerische Lösung der Flachwassergleichungen bei kleinen Froudezahlen hergeleitet. Dies beinhaltet zwei verschiedene Erweiterungen des oben genannten Projektionsverfahrens durch Berücksichtigung der lokalen Zeitableitungen des Höhenfeldes. Im Mehrskalenverfahren werden diese beiden Varianten kombiniert. Dies führt zur Lösung eines Multilevel-Helmholtzproblems analog zum linearen Fall. Das Verfahren wird für den eindimensionalen Fall implementiert, und seine Konvergenzeigenschaften werden anhand verschiedener Testfälle untersucht. Außerdem werden die Balancierungseigenschaften bzgl. eines ruhenden Sees getestet und das asymptotische Regime schneller Schwerewellen, die über kurzwellige Bodentopographie wandern, betrachtet. Die numerischen Ergebnisse des Verfahrens lassen darauf schließen, dass dieses Regime korrekt reproduziert wird, und es sich somit um ein sogenanntes asymptotisch adaptives numerisches Verfahren handelt.

A.5 Lebenslauf

Die online veröffentlichte Version enthält keinen Lebenslauf.

Bibliography

Almgren, A. S., J. B. Bell, and W. Y. Crutchfield (2000), Approximate Projection Methods: Part I. Inviscid Analysis, *SIAM Journal on Scientific Computing* **22**, 1139–1159, DOI: `10.1137/S1064827599357024`.

Asselin, R. (1972), Frequency Filter for Time Integrations, *Monthly Weather Review* **100**, 487–490, DOI: `10.1175/1520-0493(1972)100<0487:FFFTI>2.3.CO;2`.

Audusse, E., F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame (2004), A Fast and Stable Well-Balanced Scheme with Hydrostatic Reconstruction for Shallow Water Flows, *SIAM Journal on Scientific Computing* **25**, 2050–2065, DOI: `10.1137/S1064827503431090`.

Audusse, E., R. Klein, D. D. Nguyen, and S. Vater (2011), Preservation of the Discrete Geostrophic Equilibrium in Shallow Water Flows, in: J. Fořt, J. Fürst, J. Halama, R. Herbin, and F. Hubert (eds.), *Finite Volumes for Complex Applications VI – Problems & Perspectives*, Springer Berlin / Heidelberg, vol. 4 of *Springer Proceedings in Mathematics*, 59–67, DOI: `10.1007/978-3-642-20671-9_7`.

Ball, F. K. (1965), The effect of rotation on the simpler modes of motion of a liquid in an elliptic paraboloid, *Journal of Fluid Mechanics* **22**, 529–545, DOI: `10.1017/S0022112065000952`.

Bannon, P. R. (1996), On the Anelastic Approximation for a Compressible Atmosphere, *Journal of the Atmospheric Sciences* **53**, 3618–3628, DOI: `10.1175/1520-0469(1996)053<3618:OTAAFA>2.0.CO;2`.

Barreiro, A. K., and O. Bühler (2008), Longshore current dislocation on barred beaches, *Journal of Geophysical Research* **113**, C12004, DOI: `10.1029/2007JC004661`.

Bell, J. B., P. Colella, and H. M. Glaz (1989), A second-order projection method for the incompressible Navier-Stokes equations, *Journal of Computational Physics* **85**, 257–283, DOI: `10.1016/0021-9991(89)90151-4`.

Bell, J. B., and D. L. Marcus (1992), A second-order projection method for variable-density flows, *Journal of Computational Physics* **101**, 334–348, DOI: `10.1016/0021-9991(92)90011-M`.

Benoit, R., M. Desgagné, P. Pellerin, S. Pellerin, Y. Chartier, and S. Desjardins (1997), The Canadian MC2: A Semi-Lagrangian, Semi-Implicit Wideband Atmospheric Model Suited for Finescale Process Studies and Simulation, *Monthly Weather Review* **125**, 2382–2415, DOI: `10.1175/1520-0493(1997)125<2382:TCMASL>2.0.CO;2`.

Botta, N., R. Klein, S. Langenberg, and S. Lützenkirchen (2004), Well balanced finite volume methods for nearly hydrostatic flows, *Journal of Computational Physics* **196**, 539–565, DOI: `10.1016/j.jcp.2003.11.008`.

Bouchut, F. (1994), On zero pressure gas dynamics, in: B. Perthame (ed.), *Advances in kinetic theory and computing: selected papers*, World Scientific Publishing, vol. 22 of *Series on Advances in Mathematics for Applied Sciences*, 171–190.

Bouchut, F., S. Jin, and X. Li (2003), Numerical Approximations of Pressureless and Isothermal Gas Dynamics, *SIAM Journal on Numerical Analysis* **41**, 135–158, DOI: `10.1137/S0036142901398040`.

Bresch, D., R. Klein, and C. Lucas (2011), Multiscale Analyses for the Shallow Water Equations, in: E. Krause, Y. Shokin, M. Resch, D. Kröner, and N. Shokina (eds.), *Computational Science and High Performance Computing IV*, Springer Berlin / Heidelberg, vol. 115 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, 149–164, DOI: `10.1007/978-3-642-17770-5_12`.

Bühler, O., and T. E. Jacobson (2001), Wave-driven currents and vortex dynamics on barred beaches, *Journal of Fluid Mechanics* **449**, 313–339, DOI: `10.1017/S0022112001006322`.

Casulli, V., and D. Greenspan (1984), Pressure method for the numerical solution of transient, compressible fluid flows, *International Journal for Numerical Methods in Fluids* **4**, 1001–1012, DOI: `10.1002/fld.1650041102`.

Chorin, A. J. (1967), A numerical method for solving incompressible viscous flow problems, *Journal of Computational Physics* **2**, 12–26, DOI: `10.1016/0021-9991(67)90037-X`.

Chorin, A. J. (1968), Numerical Solution of the Navier-Stokes Equations, *Mathematics of Computation* **22**, 745–762, DOI: `10.1090/S0025-5718-1968-0242392-2`.

Cordier, F., P. Degond, and A. Kumbaro (2012), An Asymptotic-Preserving all-speed scheme for the Euler and Navier-Stokes equations, *Journal of Computational Physics* **231**, 5685–5704, DOI: `10.1016/j.jcp.2012.04.025`.

Courant, R., K. O. Friedrichs, and H. Lewy (1928), Über die partiellen Differenzengleichungen der mathematischen Physik, *Mathematische Annalen* **100**, 32–74, DOI: `10.1007/BF01448839`.

Currò, C. (1989), Some new exact solutions to the nonlinear shallow-water wave equations via group analysis, *Meccanica* **24**, 26–35, DOI: `10.1007/BF01576000`.

Davies, T., A. Staniforth, N. Wood, and J. Thuburn (2003), Validity of anelastic and other equation sets as inferred from normal-mode analysis, *Quarterly Journal of the Royal Meteorological Society* **129**, 2761–2775, DOI: `10.1256/qj.02.1951`.

Demirdžić, I., Ž. Lilek, and M. Perić (1993), A collocated finite volume method for predicting flows at all speeds, *International Journal for Numerical Methods in Fluids* **16**, 1029–1050, DOI: `10.1002/fld.1650161202`.

Deuflhard, P., and F. Bornemann (2002), *Scientific Computing with Ordinary Differential Equations*, vol. 42 of *Texts in Applied Mathematics*, Springer.

Dubois, T., F. Jauberteau, and R. Temam (2004), *Encyclopedia of Computational Mechanics*, John Wiley & Sons, Ltd, chap. Dynamic Multilevel Methods and Turbulence, DOI: `10.1002/0470091355.ecm059`.

Dubois, T., F. Jauberteau, R. Temam, and J. Tribbia (2005), Multilevel schemes for the shallow water equations, *Journal of Computational Physics* **207**, 660–694, DOI: `10.1016/j.jcp.2005.01.025`.

Durran, D. R. (1989), Improving the Anelastic Approximation, *Journal of the Atmospheric Sciences* **46**, 1453–1461, DOI: `10.1175/1520-0469(1989)046<1453:ITAA>2.0.CO;2`.

Durran, D. R. (2010), *Numerical Methods for Fluid Dynamics: With Applications to Geophysics*, no. 32 in Texts in Applied Mathematics, Springer, 2nd edn..

Durran, D. R., and P. N. Blossey (2012), Implicit–Explicit Multistep Methods for Fast-Wave–Slow-Wave Problems, *Monthly Weather Review* **140**, 1307–1325, DOI: `10.1175/MWR-D-11-00088.1`.

Ebin, D. G. (1982), Motion of slightly compressible fluids in a bounded domain. I, *Communications on Pure and Applied Mathematics* **35**, 451–485, DOI: `10.1002/cpa.3160350402`.

Einfeldt, B. (1988), On Godunov-Type Methods for Gas Dynamics, *SIAM Journal on Numerical Analysis* **25**, 294–318, DOI: `10.1137/0725021`.

Erbes, G. (1993), A semi-Lagrangian method of characteristics for the shallow-water equations, *Monthly Weather Review* **121**, 3443–3452, DOI: `10.1175/1520-0493(1993)121 <3443:ASLMOC>2.0.CO;2`.

Evans, K. J., M. A. Taylor, and J. B. Drake (2010), Accuracy Analysis of a Spectral Element Atmospheric Model Using a Fully Implicit Solution Framework, *Monthly Weather Review* **138**, 3333–3341, DOI: `10.1175/2010MWR3288.1`.

Fukushima, T. (1999), Super Implicit Multistep Methods, in: H. Umehara (ed.), *31st Symposium on Celestial Mechanics*, Kashima Space Research Center, Ibaraki, Japan, 343–366.

Gadd, A. J. (1978), A split explicit integration scheme for numerical weather prediction, *Quarterly Journal of the Royal Meteorological Society* **104**, 569–582, DOI: `10.1002/qj.49710444103`.

Geratz, K. J. (1997), *Erweiterung eines Godunov-Typ-Verfahrens für zwei-dimensionale kompressible Strömungen auf die Fälle kleiner und verschwindender Machzahl*, Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen.

Giraldo, F. X. (2005), Semi-implicit time-integrators for a scalable spectral element atmospheric model, *Quarterly Journal of the Royal Meteorological Society* **131**, 2431–2454, DOI: `10.1256/qj.03.218`.

Gosse, L., and A.-Y. LeRoux (1996), A well-balanced scheme designed for inhomogeneous scalar conservation laws, *Comptes Rendus de l'Academie des Sciences-Serie I-Mathematique* **323**, 543–546.

Gottlieb, S., C.-W. Shu, and E. Tadmor (2001), Strong Stability-Preserving High-Order Time Discretization Methods, *SIAM Review* **43**, 89–112, DOI: `10.1137/S003614450036757X`.

Grabowski, W. W. (1998), Toward Cloud Resolving Modeling of Large-Scale Tropical Circulations: A Simple Cloud Microphysics Parameterization, *Journal of the Atmospheric Sciences* **55**, 3283–3298, DOI: `10.1175/1520-0469(1998)055<3283:TCRMOL>2.0.CO;2`.

Greenberg, J. M., and A.-Y. LeRoux (1996), A Well-Balanced Scheme for the Numerical Processing of Source Terms in Hyperbolic Equations, *SIAM Journal on Numerical Analysis* **33**, 1–16.

Greenspan, H. P. (1968), *The theory of rotating fluids*, Cambridge Monographs on Mechanics and Applied Mathematics, Cambridge University Press, Cambridge.

Grinstein, F. F., L. G. Margolin, and W. J. Rider (2007), *Implicit Large Eddy Simulation*, Cambridge University Press, Cambridge, DOI: `10.1017/CBO9780511618604`.

Guermond, J.-L., P. Minev, and J. Shen (2006), An overview of projection methods for incompressible flows, *Computer Methods in Applied Mechanics and Engineering* **195**, 6011–6045, DOI: `10.1016/j.cma.2005.10.010`.

Hairer, E., C. Lubich, and G. Wanner (2006), *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, vol. 31 of *Springer Series in Computational Mathematics*, Springer, 2nd edn..

Harlow, F. H., and J. E. Welch (1965), Numerical Calculation of Time-Dependent Viscous Incompressible Flow of Fluid with Free Surface, *The Physics of Fluids* **8**, 2182–2189, DOI: `10.1063/1.1761178`.

Hickel, S., N. A. Adams, and J. A. Domaradzki (2006), An adaptive local deconvolution method for implicit LES, *Journal of Computational Physics* **213**, 413–436, DOI: `10.1016/j.jcp.2005.08.017`.

Iacono, R. (2005), Analytic solutions to the shallow water equations, *Physical Review E* **72**, 017302, DOI: `10.1103/PhysRevE.72.017302`.

Jebens, S., O. Knoth, and R. Weiner (2009), Explicit Two-Step Peer Methods for the Compressible Euler Equations, *Monthly Weather Review* **137**, 2380–2392, DOI: `10.1175/2008MWR2671.1`.

Jebens, S., O. Knoth, and R. Weiner (2012), Linearly implicit peer methods for the compressible Euler equations, *Applied Numerical Mathematics* **62**, 1380–1392, DOI: `10.1016/j.apnum.2012.06.013`.

Johnson, E. R., A. K. Hinds, and N. R. McDonald (2005), Steadily translating vortices near step topography, *Physics of Fluids* **17**, 056601, DOI: `10.1063/1.1890427`.

Kadioglu, S. Y., R. Klein, and M. L. Minion (2008), A fourth-order auxiliary variable projection method for zero-Mach number gas dynamics, *Journal of Computational Physics* **227**, 2012–2043, DOI: `10.1016/j.jcp.2007.10.008`.

Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge, DOI: `10.1017/CBO9780511802270`.

Karki, K., and S. Patankar (1988), Calculation procedure for viscous incompressible flows in complex geometries, *Numerical Heat Transfer* **14**, 295–307, DOI: `10.1080/10407788808913645`.

Kevorkian, J., and J. D. Cole (1996), *Multiple Scale and Singular Perturbation Methods*, vol. 114 of *Applied Mathematical Sciences*, Springer, New York.

Klainerman, S., and A. Majda (1981), Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids, *Communications in Pure Applied Mathematics* **34**, 481–524, DOI: `10.1002/cpa.3160340405`.

Klein, R. (1995), Semi-Implicit Extension of a Godunov-Type Scheme Based on Low Mach Number Asymptotics I: One-Dimensional Flow, *Journal of Computational Physics* **121**, 213–237, DOI: `10.1016/S0021-9991(95)90034-9`.

Klein, R. (2000), Asymptotic Analyses for Atmospheric Flows and the Construction of Asymptotically Adaptive Numerical Methods, *Zeitschrift für Angewandte Mathematik und Mechanik* **80**, 765–777.

Klein, R. (2004), An Applied Mathematical View of Meteorological Modelling, in: J. M. Hill and R. Moore (eds.), *Applied Mathematics Entering the 21st Century: Invited talks from the ICIAM 2003 Congress*, vol. 116 of *Proceedings in Applied Mathematics*, 227–269.

Klein, R. (2009), Asymptotics, structure, and integration of sound-proof atmospheric flow equations, *Theoretical and Computational Fluid Dynamics* **23**, 161–195, DOI: `10.1007/s00162-009-0104-y`.

Klein, R. (2010), Scale-Dependent Models for Atmospheric Flows, *Annual Review of Fluid Mechanics* **42**, 249–274, DOI: `10.1146/annurev-fluid-121108-145537`.

Klein, R., U. Achatz, D. Bresch, O. M. Knio, and P. K. Smolarkiewicz (2010), Regime of Validity of Soundproof Atmospheric Flow Models, *Journal of the Atmospheric Sciences* **67**, 3226–3237, DOI: `10.1175/2010JAS3490.1`.

Klein, R., N. Botta, T. Schneider, C.-D. Munz, S. Roller, A. Meister, L. Hoffmann, and T. Sonar (2001), Asymptotic adaptive methods for multi-scale problems in fluid mechanics, *Journal of Engineering Mathematics* **39**, 261–343, DOI: `10.1023/A:1004844002437`.

Klein, R., S. Vater, E. Päschke, and D. Ruprecht (2011), Multiple scales methods in meteorology, in: H. Steinrück (ed.), *Asymptotic Methods in Fluid Mechanics: Survey and Recent Advances*, Springer, vol. 523 of *CISM International Centre for Mechanical Sciences*, 127–196, URL `http://www.springer.com/materials/mechanics/book/978-3-7091-0407-1`.

Klemp, J. B., and R. B. Wilhelmson (1978), The Simulation of Three-Dimensional Convective Storm Dynamics, *Journal of the Atmospheric Sciences* **35**, 1070–1096, DOI: `10.1175/1520-0469(1978)035<1070:TSOTDC>2.0.CO;2`.

Knoth, O. (2006), Compressible atmospheric modeling at all scales, *Oberwolfach Reports* **3**, 2374–2377, DOI: `10.4171/OWR/2006/39`, workshop: Mathematical Theory and Modelling in Atmosphere-Ocean Science.

Kwizak, M., and A. J. Robert (1971), A Semi-Implicit Scheme for Grid Point Atmospheric Models of the Primitive Equations, *Monthly Weather Review* **99**, 32–36, DOI: `10.1175/1520-0493(1971)099<0032:ASSFGP>2.3.CO;2`.

Le Maître, O., J. Levin, M. Iskandarani, and O. M. Knio (2001), A Multiscale Pressure Splitting of the Shallow-Water Equations I. Formulation and 1D Tests, *Journal of Computational Physics* **166**, 116–151, DOI: `10.1006/jcph.2001.6647`.

LeVeque, R. J. (1998), Balancing Source Terms and Flux Gradients in High-Resolution Godunov Methods: The Quasi-Steady Wave-Propagation Algorithm, *Journal of Computational Physics* **146**, 346–365, DOI: `10.1006/jcph.1998.6058`.

LeVeque, R. J. (2002), *Finite Volume Methods for Hyperbolic Problems*, vol. 31 of *Cambridge Texts in Applied Mathematics*, Cambridge University Press, Cambridge.

LeVeque, R. J. (2004), The dynamics of pressureless dust clouds and delta waves, *Journal of Hyperbolic Differential Equations* **1**, 315–327, DOI: `10.1142/S0219891604000135`.

Levermore, C. D., M. Oliver, and E. S. Titi (1996), Global well-posedness for the lake equations, *Physica D: Nonlinear Phenomena* **98**, 492–509, DOI: `10.1016/0167-2789(96)00108-X`.

Lipps, F. B., and R. S. Hemler (1982), A Scale Analysis of Deep Moist Convection and Some Related Numerical Calculations, *Journal of the Atmospheric Sciences* **39**, 2192–2210, DOI: `10.1175/1520-0469(1982)039<2192:ASAODM>2.0.CO;2`.

Majda, A. J. (2003), *Introduction to PDEs and Waves for the Atmosphere and Ocean*, vol. 9 of *Courant Lecture Notes in Mathematics*, AMS, New York.

Majda, A. J., and R. Klein (2003), Systematic Multiscale Models for the Tropics, *Journal of the Atmospheric Sciences* **60**, 393–408, DOI: `10.1175/1520-0469(2003)060<0393:SMMFTT>2.0.CO;2`.

Munz, C.-D., S. Roller, R. Klein, and K. J. Geratz (2003), The extension of incompressible flow solvers to the weakly compressible regime, *Computers & Fluids* **32**, 173–196.

Nerinckx, K., J. Vierendeels, and E. Dick (2005), Mach-uniformity through the coupled pressure and temperature correction algorithm, *Journal of Computational Physics* **206**, 597–623, DOI: `10.1016/j.jcp.2004.12.019`.

Ogura, Y., and N. A. Phillips (1962), Scale Analysis of Deep and Shallow Convection in the Atmosphere, *Journal of the Atmospheric Sciences* **19**, 173–179, DOI: `10.1175/1520-0469(1962)019<0173:SAODAS>2.0.CO;2`.

Ohfuchi, W., H. Nakamura, M. Yoshioka, T. Enomoto, K. Takaya, X. Peng, S. Yamane, T. Nishimura, Y. Kurihara, and K. Ninomiya (2004), 10-km mesh meso-scale resolving simulations of the global atmosphere on the Earth Simulator: Preliminary outcomes of AFES (AGCM for the Earth Simulator), *Journal of the Earth Simulator* **1**, 8–34.

Oliver, M. (1997), Justification of the Shallow-Water Limit for a Rigid-Lid Flow with Bottom Topography, *Theoretical and Computational Fluid Dynamics* **9**, 311–324, DOI: `10.1007/s001620050047`.

Osher, S. (1985), Convergence of generalized MUSCL schemes, *SIAM Journal on Numerical Analysis* **22**, 947–961, DOI: `10.1137/0722057`.

Park, J.-H. (2003), *Ein konservatives MPV-Verfahren zur Simulation der Strömungen in allen Machzahlbereichen*, Dissertation, Universität Stuttgart, URL `http://elib.uni-stuttgart.de/opus/volltexte/2003/1530/`.

Patankar, S., and D. Spalding (1972), A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows, *International Journal of Heat and Mass Transfer* **15**, 1787–1806, DOI: `10.1016/0017-9310(72)90054-3`.

Patnaik, G., R. Guirguis, J. Boris, and E. Oran (1987), A barely implicit correction for flux-corrected transport, *Journal of Computational Physics* **71**, 1–20, DOI: `10.1016/0021-9991(87)90016-7`.

Pedlosky, J. (1987), *Geophysical Fluid Dynamics*, Springer, New York, 2nd edn..

Reisner, J., V. Mousseau, and D. Knoll (2001), Application of the Newton–Krylov Method to Geophysical Flows, *Monthly Weather Review* **129**, 2404–2415, DOI: `10.1175/1520-0493(2001)129<2404:AOTNKM>2.0.CO;2`.

Reisner, J., A. Wyszogrodzki, V. Mousseau, and D. Knoll (2003), An efficient physics-based preconditioner for the fully implicit solution of small-scale thermally driven atmospheric flows, *Journal of Computational Physics* **189**, 30–44, DOI: `10.1016/S0021-9991(03)00198-0`.

Reisner, J. M., A. Mousseau, A. Wyszogrodzki, and D. A. Knoll (2005), An Implicitly Balanced Hurricane Model with Physics-Based Preconditioning, *Monthly Weather Review* **133**, 1003–1022, DOI: `10.1175/MWR2901.1`.

Rieper, F. (2008), *On the behaviour of numerical schemes in the low Mach number regime*, Dissertation, Brandenburg Technical University, URL `http://opus.kobv.de/btu/volltexte/2008/574/`.

Rieper, F. (2011), A low-Mach number fix for Roe's approximate Riemann solver, *Journal of Computational Physics* **230**, 5263–5287, DOI: `10.1016/j.jcp.2011.03.025`.

Robert, A. J. (1966), The Integration of a Low Order Spectral Form of the Primitive Meteorological Equations, *Journal of the Meteorological Society of Japan* **44**, 237–245.

Robert, A. J. (1969), The integration of a spectral model of the atmosphere by the implicit method, in: *Proceedings of the WMO/IUGG Symposium on Numerical Weather Prediction, Tokyo, November 26-December 4, 1968.*, Meterological Society of Japan, Tokyo, vol. 7, 19–24.

Robert, A. J. (1982), A Semi-Lagrangian and Semi-Implicit Numerical Integration Scheme for the Primitive Meteorological Equations, *Journal of the Meteorological Society of Japan* **60**, 319–325.

Schneider, T., N. Botta, K. J. Geratz, and R. Klein (1999), Extension of Finite Volume Compressible Flow Solvers to Multi-dimensional, Variable Density Zero Mach Number Flows, *Journal of Computational Physics* **155**, 248–286, DOI: `10.1006/jcph.1999.6327`.

Schneider, W. (1978), *Mathematische Methoden der Strömungsmechanik*, Vieweg & Sohn Verlagsgesellschaft, Braunschweig.

Schochet, S. (1988), Asymptotics for symmetric hyperbolic systems with a large parameter, *Journal of Differential Equations* **75**, 1–27, DOI: `10.1016/0022-0396(88)90126-X`.

Schochet, S. (2005), The mathematical theory of low Mach number flows, *RAIRO Modélisation Mathématique et Analyse Numérique* **39**, 441–458, DOI: `10.1051/m2an:2005017`.

Sesterhenn, J., B. Müller, and H. Thomann (1999), On the Cancellation Problem in Calculating Compressible Low Mach Number Flows, *Journal of Computational Physics* **151**, 597–615, DOI: `10.1006/jcph.1999.6211`.

Shapiro, A. (1996), Nonlinear shallow-water oscillations in a parabolic channel: exact solutions and trajectory analyses, *Journal of Fluid Mechanics* **318**, 49–76, DOI: `10.1017/S0022112096007021`.

Shu, C.-W., and S. Osher (1988), Efficient Implementation of Essentially Non-oscillatory Shock-Capturing Schemes, *Journal of Computational Physics* **77**, 439–471, DOI: `10.1016/0021-9991(88)90177-5`.

Skamarock, W. C., and J. B. Klemp (1992), The Stability of Time-Split Numerical Methods for the Hydrostatic and the Nonhydrostatic Elastic Equations, *Monthly Weather Review* **120**, 2109–2127, DOI: `10.1175/1520-0493(1992)120<2109:TSOTSN>2.0.CO;2`.

Smolarkiewicz, P. K., and A. Dörnbrack (2008), Conservative integrals of adiabatic Durran's equations, *International Journal for Numerical Methods in Fluids* **56**, 1513–1519, DOI: `10.1002/fld.1601`.

St-Cyr, A., and D. Neckels (2009), A Fully Implicit Jacobian-Free High-Order Discontinuous Galerkin Mesoscale Flow Solver, in: G. Allen, J. Nabrzyski, E. Seidel, G. van Albada, J. Dongarra, and P. Sloot (eds.), *Computational Science – ICCS 2009*, Springer Berlin / Heidelberg, vol. 5545 of *Lecture Notes in Computer Science*, 243–252, DOI: `10.1007/978-3-642-01973-9_28`.

Süli, E. (1991), Convergence of finite volume schemes for Poisson's equation on nonuniform meshes, *SIAM Journal on Numerical Analysis* **28**, 1419–1430, DOI: `10.1137/0728073`.

Temam, R. (1968), Une méthode d'approximation de la solution des équations de Navier-Stokes, *Bulletin de la Société Mathématique de France* **96**, 115–152, URL `http://www.numdam.org/item?id=BSMF_1968__96__115_0`.

Thacker, W. C. (1981), Some exact solutions to the nonlinear shallow-water wave equations, *Journal of Fluid Mechanics* **107**, 499–508, DOI: `10.1017/S0022112081001882`.

Trottenberg, U., C. Oosterlee, and A. Schüller (2001), *Multigrid*, Academic Press.

Ullrich, P., and C. Jablonowski (2012), Operator-Split Runge–Kutta–Rosenbrock Methods for Nonhydrostatic Atmospheric Models, *Monthly Weather Review* **140**, 1257–1284, DOI: `10.1175/MWR-D-10-05073.1`.

Vallis, G. K. (2006), *Atmospheric and Oceanic Fluid Dynamics*, Cambridge University Press, Cambridge.

van der Heul, D., C. Vuik, and P. Wesseling (2003), A conservative pressure-correction method for flow at all speeds, *Computers & Fluids* **32**, 1113–1132, DOI: `10.1016/S0045-7930(02)00086-5`.

van der Vorst, H. A. (1992), Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, *SIAM Journal on Scientific and Statistical Computing* **13**, 631–644, DOI: `10.1137/0913035`.

van Kan, J. (1986), A Second-Order Accurate Pressure-Correction Scheme for Viscous Incompressible Flow, *SIAM Journal on Scientific and Statistical Computing* **7**, 870–891, DOI: `10.1137/0907059`.

van Leer, B. (1979), Towards the Ultimate Conservative Difference Scheme. V. A Second-Order Sequel to Godunov's Method, *Journal of Computational Physics* **32**, 101–136, DOI: `10.1016/0021-9991(79)90145-1`.

Vater, S. (2005), A New Projection Method for the Zero Froude Number Shallow Water Equations, PIK Report 97, Potsdam Institute for Climate Impact Research, URL `http://www.pik-potsdam.de/research/publications/pikreports/.files/pr97.pdf`.

Vater, S., and R. Klein (2009), Stability of a Cartesian grid projection method for zero Froude number shallow water flows, *Numerische Mathematik* **113**, 123–161, DOI: `10.1007/s00211-009-0224-8`.

Vater, S., R. Klein, and O. M. Knio (2011), A Scale-selective Multilevel Method for Long-Wave Linear Acoustics, *Acta Geophysica* **59**, 1076–1108, DOI: `10.2478/s11600-011-0037-x`.

Welch, J. E., F. H. Harlow, J. P. Shannon, and B. J. Daly (1965), The MAC method - a computing technique for solving viscous, incompressible, transient fluid-flow problems involving free surfaces, Tech. Rep. LA-3425, Los Alamos Scientific Laboratory.

Wensch, J., O. Knoth, and A. Galant (2009), Multirate infinitesimal step methods for atmospheric flow simulation, *BIT Numerical Mathematics* **49**, 449–473, DOI: `10.1007/s10543-009-0222-3`.

Wicker, L. J., and W. C. Skamarock (1998), A Time-Splitting Scheme for the Elastic Equations Incorporating Second-Order Runge–Kutta Time Differencing, *Monthly Weather Review* **126**, 1992–1999, DOI: `10.1175/1520-0493(1998)126<1992:ATSSFT>2.0.CO;2`.

Wicker, L. J., and W. C. Skamarock (2002), Time-Splitting Methods for Elastic Models Using Forward Time Schemes, *Monthly Weather Review* **130**, 2088–2097, DOI: `10.1175/1520-0493(2002)130<2088:TSMFEM>2.0.CO;2`.