8-31-2020

Efficient time-stepping approaches for the dispersive shallow water equations

Linwan Feng
New Jersey Institute of Technology

Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

ABSTRACT

EFFICIENT TIME-STEPPING APPROACHES FOR THE DISPERSIVE SHALLOW WATER EQUATIONS

by
Linwan Feng

This dissertation focuses on developing efficient and stable (high order) time-stepping strategies for the dispersive shallow water equations (DSWE) with variable bathymetry. The DSWE extends the regular shallow water equations to include dispersive effects. Dispersion is physically important and can maintain the shape of a wave that would otherwise form a shock in the shallow water system.

In some cases, the DSWE may be simplified when the bathymetry length scales are small (or large) in relation to other length scales in the shallow water system. These simplified DSWE models, which are related to the full DSWEs, are also considered in this thesis as well.

Incorporating dispersive effects creates added difficulties when devising efficient high order time-stepping methods. Time-stepping the DSWE is difficult as the equations may be stiff as well as non-local and nonlinear in the time derivative of the velocity variables. In this dissertation, the DSWE are recast as an evolution equation in time, plus an elliptic constraint equation in space. When discretized (in space), the system of equations takes the form of an (index-1) differential-algebraic equation (DAE). Here the algebraic equation in the DAE captures dispersive effects and consists of a quasi-linear or semi-linear operator. Two strategies are examined to time-step the DSWE in constraint form — the key novelty is on solving the resulting DAE while avoiding complex nonlinear solutions to the algebraic equations:
(i) preconditioned iterative methods are devised to invert the (semi-linear) operator;
(ii) semi-implicit time-stepping (ImEx) methods are devised that bypass a full inversion of the quasi/semi-linear operator. Guaranteeing stability for the semi-

implicit approach is a nontrivial issue due to the fact that certain stiff terms in the equation are treated explicitly. A stability theory is provided which outlines how to choose the semi-implicit terms in such a way to guarantee numerical (zero) stability.

EFFICIENT TIME-STEPPING APPROACHES FOR THE DISPERSIVE SHALLOW WATER EQUATIONS

by
Linwan Feng

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
and Rutgers, The State University of New Jersey - Newark
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Mathematical Sciences

Department of Mathematical Sciences, NJIT
Department of Mathematics and Computer Science, Rutgers-Newark

August 2020

APPROVAL PAGE

EFFICIENT TIME-STEPPING APPROACHES FOR THE DISPERSIVE SHALLOW WATER EQUATIONS

Linwan Feng

Dr. David Shirokoff, Dissertation Advisor — Date
Assistant Professor, Department of Mathematical Sciences, NJIT

Dr. Wooyoung Choi, Dissertation Co - Advisor — Date
Professor, Department of Mathematical Sciences, NJIT

Dr. Richard Moore, Committee Member — Date
Director of Programs and Services, Society for Industrial and Applied Mathematics

---

Dr. Yassine Boubendir, Committee Member — Date
Professor, Department of Mathematical Sciences, NJIT

Dr. Benjamin Seibold, Committee Member — Date
Associate Professor, Department of Mathematics, Temple University

BIOGRAPHICAL SKETCH

Author: Linwan Feng

Degree: Doctor of Philosophy

Date: August 2020

Undergraduate and Graduate Education:

- Doctor of Philosophy in Mathematical Sciences,
  New Jersey Institute of Technology, Newark, NJ, 2020
- Master of Science in Applied Mathematics,
  New Jersey Institute of Technology, Newark, NJ, 2014
- Bachelor of Science in Mathematics and Applied Mathematics,
  Northwestern Polytechnical University, Xi'an, Shaanxi, China, 2012

Major: Mathematical Sciences

Presentations and Publications:

L. Feng, D. Shirokoff and W.Choi "Efficient Semi-Implicit Time-Stepping Schemes for the Dispersive Shallow Water Equations," In preparation (2020).

L Feng, "Numerical Methods for Dispersive Shallow Water Equations," *2019 Dana Knox Student Research Showcase*, NJIT, 2019.

L Feng, "Numerical Methods for Dispersive Shallow Water Equations," *2018 Mid-Atlantic Numerical Analysis Day*, Temple University, 2018.

L Feng, "Numerical Methods for Dispersive Shallow Water Equations," *2018 American Mathematical Society Fall Eastern Sectional Meeting*, University of Delaware, 2018.

Dedicated to my mom and dad.

Love you so much and you are the best.

给我的父母

是你们让我有自由的成长环境。

还有家人们一直以来的关心与爱护。

爱你们 ♡

ACKNOWLEDGMENT

I would like to express my deep gratitude to Professor David Shirokoff and Professor Wooyoung Choi, my research supervisors, for their patient guidance, enthusiastic encouragement and useful critiques of this research work.

I also appreciate Dr. Richard Moore, Dr. Yassine Boubendir and Dr. Benjamin Seibold for their kindness of being my committee members.

I would like to thank Dr. Micheal Siegel, for his advice and assistance in keeping my progress on schedule.

My grateful thanks are also extended to Department of Mathematical Sciences of NJIT and the National Science Foundation Grant No. DMS-1719693 for supporting during my years of research.

Finally, I wish to thank my parents and my family for their support and encouragement throughout my study.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This dissertation devises stable and efficient time-stepping methods for dispersive shallow water (DSW) equations [10] under various bathymetry. With bathymetry included in the system, the dispersive terms contain mixed time and space derivatives, which make the system numerically complicated to solve. Throughout this research, we always use a pseudo-spectral method [23] for spatial discretizations to give us high spatial accuracy. For smooth solutions on periodic domains, spectral methods (almost always) yield higher accuracy than finite element/volume methods.

Nonlinear long water waves in shallow water have been extensively studied for a wide range of their applications in coastal oceans. After imposing the balance between nonlinearity and dispersion under the small amplitude assumption, a number of weakly nonlinear and weakly dispersive models, often referred to as Boussinesq-type equations, have been proposed. These models typically contain the leading-order nonlinear and dispersive terms for small amplitude long waves and describe the two-dimensional evolution of the surface elevation and the depth-averaged horizontal velocities. See the lists of references in Whitham [25], Mei [12], and Wu [27]. In particular, the propagation and transformation of solitary waves over bottom topography has been a main research topic for Tsunami modeling [13]. The Boussinesq-type models have been often solved numerically using finite difference schemes, e.g., by Peregrine [17], Nwogu [15], and many others, and it has been found that the numerical solutions agree well with laboratory and field experiments [22, 11] when the wave amplitudes are relatively small compared with the water depth.

As the wave amplitude increases, the weakly nonlinear models are no longer valid and higher-order nonlinear effects have to be taken into account. In recent

years, much attention has been paid to strongly nonlinear long wave models, for which no assumption on the wave amplitude is made while the wavelength is still assumed to be much greater than the water depth. A strongly nonlinear model with the leading-order dispersive terms has been suggested for the depth-averaged velocity by various researchers, including Rayleigh [8] and Serre [20] for steady waves and by Su & Gardner [21] and Green & Naghdi [6] for unsteady waves. It has been shown by Li *et al.* [10] that the model is reliable for the evolution of solitary waves with relatively large amplitudes in comparison with the numerical solutions of the fully nonlinear Euler equations. Instead of the depth-averaged velocity, Wei *et al.* [5] proposed a model for the bottom velocity that is asymptotically equivalent to the strongly nonlinear model mentioned previously, and validated the numerical solutions with their own experiments. A similar approach has been also extended to the two-layer problem [14, 2]. The leading-order strongly nonlinear long wave model has been further extended to an arbitrary order of nonlinearity by Wu [26], Agnon *et al.* [28] and Madsen *et al.* [16] for various dependent variables although the system becomes ill-posed under even-order approximation for the depth-average velocity and odd-order for the surface velocity [1].

When the dispersive terms are completely neglected, the strongly nonlinear long wave model reduces to the non-dispersive shallow water equations, which have been widely used for large-scale ocean circulation modeling (although the circulation models solve multi-layer versions). To correctly describe smaller-scale motions in the ocean, asymptotically-consistent dispersive terms with high-order spatial derivatives need to be included, but have been neglected as the computational cost to solve the dispersive shallow water (DSW) equations is too high. It should be remarked that, even for the non-dispersive case, the ocean circulation models usually adopt relatively low-order finite difference schemes for both spatial and temporal discretizations. Even for regional ocean modeling, this is not so surprising as the size of computational

domain, and, therefore, the number of grid points are overwhelmingly large. Thus, it is desirable to develop an efficient numerical scheme to solve the DSW equations or the strongly nonlinear long wave model.

This dissertation aims to develop a reliable numerical scheme to solve the DSWE with a variable bottom, for possible oceanic applications. More specifically, after adopting a pseudo-spectral scheme for spatial discretization, we explore an efficient and stable time-stepping scheme. After the spatial discretization, we recast the dispersive shallow water equation into a differential-algebraic form and then derive an efficient time-stepping method. For the time-stepping method, we pursue two strategies: (i) Invert the quasi/semi-linear operator with matrix-free iterative method that can be interpreted as a preconditioned gradient descent [4, 3], and (ii) introduce a new semi-implicit time-stepping method [18, 19]. In this research, the stability [9] and efficiency of these two methods is the main focus.

CHAPTER 2

DISPERSIVE SHALLOW WATER EQUATIONS

In this chapter, the dispersive shallow water (DSW) equations are introduced. Section 2.1 focuses on the derivation of the DSWE model and provides the basic background definitions and notations as well. Section 2.2 mainly discusses a recast version of the DSWE that, when numerically discretized, yields a differential-algebraic equation. Specifically, the DSWE can be recast into a time evolution PDE plus a spatial PDE constraint which can give an easier form to establish stability and numerical schemes. While in Section 2.3, by discussing the various approximations for small parameters, the DSWE can be reduced to simpler models.

2.1 Background Definitions and Notations

The dispersive shallow water equations are derived from the incompressible Euler equations under the long wave approximation. The three-dimensional (3D) Euler equations are given by

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} = -\frac{\nabla p}{\rho} - g, \tag{2.1}$$

$$\frac{\partial \rho}{\partial t} + \boldsymbol{u} \cdot \nabla \rho = 0, \tag{2.2}$$

$$\nabla \cdot \boldsymbol{u} = 0. \tag{2.3}$$

In equations (2.1)–(2.3), $\boldsymbol{u} = (\widehat{\boldsymbol{u}}, w)$ is the fluid velocity which $\widehat{\boldsymbol{u}} \in \mathbb{R}^D$ (where $D = 1$, or 2) is the horizontal component and w is the vertical component; $\rho(\boldsymbol{x}, z, t)$ is the fluid density; $p(\boldsymbol{x}, z, t)$ is the pressure; g is gravitational acceleration. Here $\nabla = (\nabla_{\boldsymbol{x}}, \partial_z)$ is the spatial gradient in $\mathbb{R}^{D+1}$.

From the Euler equations, after introducing the shallow water (long-wave) parameter β, the Shallow Water Equations (SWE) and Dispersive Shallow Water Equations (DSWE) can be derived, after assuming that β is small, as the first- and second-order approximations, respectively. Here the long-wave parameter β is the ratio between the mean water depth $\bar{h}$ and the wave length λ (as defined in Figure 2.1).

$$\text{Shallow water (long-wave) parameter:} \qquad \beta \equiv \frac{\bar{h}}{\lambda} \ll 1. \tag{2.4}$$

For long surface waves, in addition to β that measures dispersion, another small parameter ϵ should be introduced to measure nonlinearity. Here ϵ is the radio of the wave amplitude a to the mean water depth $\bar{h}$. In this dissertation, we are interested in a strongly nonlinear system valid for $\epsilon \equiv \frac{a}{\bar{h}} = O(1)$.

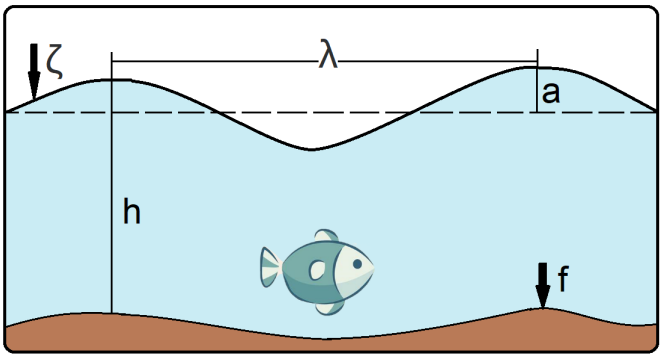


Figure 2.1 Shallow water wave system with corresponding parameters (scales may differ).

To derive the SWE and DSWE, for $\beta \ll 1$, but $\epsilon = O(1)$, from the Euler equations (2.1)–(2.3), we introduce the depth-averaged velocity $\bar{u}$ defined as

$$\bar{u}(\boldsymbol{x}, t) = \frac{1}{\zeta + h} \int_{-h}^{\zeta} u(\boldsymbol{x}, z, t) \ \mathrm{d}z. \tag{2.5}$$

In equation (2.5), $\zeta(\boldsymbol{x}, t)$ is the free surface elevation, $h(\boldsymbol{x})$ is bathymetry, and $\eta(\boldsymbol{x}, t) = h + \zeta$ is the local water depth. The SWE and DSWE can be derived by substituting the following asymptotic expansions

$$(u, \omega, p) = (u_0, \omega_0, p_0) + \beta^2(u_1, \omega_1, p_1) + O(\beta^4) \tag{2.6}$$

into the non-dimensionalized Euler equations (the detailed derivations are provided in Appendix). The leading-order approximation at $O(1)$ yields the shallow water equations (SWE).

$$\eta_t + (\eta\bar{u})_x = 0, \tag{2.7}$$

$$\bar{u}_t + \bar{u}\bar{u}_x + \eta_x = 0, \tag{2.8}$$

When truncated at $O(\beta^2)$, one can obtain the dispersive shallow water equations (DSWE), which can be written for $D = 1$ as

$$\eta_t + (\eta\bar{u})_x = 0, \tag{2.9}$$

$$\bar{u}_t + \bar{u}\bar{u}_x + \eta_x = \frac{1}{\eta}\left(\frac{\eta^3}{3}G + \frac{\eta^2}{2}F\right)_x - \left(\frac{\eta}{2}G + F\right)h_x + O(\beta^4), \tag{2.10}$$

where $G = \bar{u}_{xt} + \bar{u}\bar{u}_{xx} - \bar{u}_x^2$, $F = u_t + u(uh_x)_x$ and $\eta = \zeta + h$. While equation (2.9) implying mass conservation is exact, the momentum equation (2.9) is valid to $O(\beta^2)$. On the other hand, the shallow water equations (SWE), truncate at $O(1)$, can be obtained by neglecting the dispersive terms of $O(\beta^2)$ on the right-hand side of equation (2.10).

Notice that the nondispersive SWE will develop shocks, especially when $D = 1$. So including the right-hand side of the equation (2.10) that represents dispersive effects can regularize the SWE. Hereafter, for brevity, the bar for the depth-averaged velocity $\bar{u}$ will be dropped. With the same approach as for $D = 1$, equations (2.9)–(2.10) can be extended to the two-dimensional case:

$$\eta_t + \nabla \cdot (\eta \boldsymbol{u}) = 0 \,, \tag{2.11}$$

$$\begin{aligned}
&\underbrace{\boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla \boldsymbol{u} + \nabla \zeta}_{\text{SWE terms}} = \\
&\frac{1}{\eta} \nabla \left\{ \frac{1}{3} \eta^3 \left(\nabla \cdot \boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla \left(\nabla \cdot \boldsymbol{u} \right) - \left(\nabla \cdot \boldsymbol{u} \right)^2 \right) + \frac{1}{2} \eta^2 \left(\boldsymbol{u}_t \cdot \nabla h + \left(\boldsymbol{u} \cdot \nabla \right)^2 h \right) \right\} \\
&- \left\{ \frac{1}{2} \eta \left(\nabla \cdot \boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla \left(\nabla \cdot \boldsymbol{u} \right) - \left(\nabla \cdot \boldsymbol{u} \right)^2 \right) + \left(\boldsymbol{u}_t \cdot \nabla h + \left(\boldsymbol{u} \cdot \nabla \right)^2 h \right) \right\} \nabla h \,.
\end{aligned} \tag{2.12}$$

The boundary conditions imposed in this dissertation are doubly periodic boundary conditions, which means if $(x, y) \in \Omega = (0, L_x) \times (0, L_y)$

$$\boldsymbol{u}(x + L_x, y) = \boldsymbol{u}(x, y), \quad \boldsymbol{u}(x, y + L_y) = \boldsymbol{u}(x, y), \tag{2.13}$$

$$\eta(x + L_x, y) = \eta(x, y), \quad \eta(x, y + L_y) = \eta(x, y). \tag{2.14}$$

In this research, the main objective is to solve equations (2.11)–(2.12) using a pseudo-spectral method for spatial discretizations with an efficient time-stepping method. The main difficulty lies in the stiff terms with mixed third-order derivatives in time and space on the right-hand side of equation (2.12).

2.2 Differential-Algebraic Form of DSWE

Due to complexity in the DSWE system, to solve the equations with an efficient time-stepping method, the DSWE (2.11)–(2.12) are rewritten as a system of time

evolution PDEs plus a spatial PDE constraint. The form of differential algebraic equations (DAE) is convenient to develop a numerical scheme and test its stability.

The first step is to bring all terms involving $\boldsymbol{u}_t$ to the left-hand side in equation (2.12) to have

$$\begin{aligned}\mathcal{N}\boldsymbol{u}_t + \eta\boldsymbol{u}\cdot\nabla\boldsymbol{u} + \eta\nabla\zeta = \nabla\left\{\frac{1}{3}\eta^3\left(\boldsymbol{u}\cdot\nabla\left(\nabla\cdot\boldsymbol{u}\right) - \left(\nabla\cdot\boldsymbol{u}\right)^2\right) + \frac{1}{2}\eta^2\left(\boldsymbol{u}\cdot\nabla\right)^2 h\right\} \\ - \eta\left\{\frac{1}{2}\eta\left(\boldsymbol{u}\cdot\nabla\left(\nabla\cdot\boldsymbol{u}\right) - \left(\nabla\cdot\boldsymbol{u}\right)^2\right) + \left(\boldsymbol{u}\cdot\nabla\right)^2 h\right\}\nabla h,\end{aligned} \tag{2.15}$$

where

$$\mathcal{N} \equiv \eta\boldsymbol{I} - (\nabla\tfrac{1}{3}\eta^3\nabla\cdot) + \eta\nabla h(\nabla h^T\cdot) - \nabla\left(\tfrac{1}{2}\eta^2\nabla h^T\cdot\right) + \tfrac{1}{2}\eta^2\nabla h(\nabla\cdot), \tag{2.16}$$

with $\boldsymbol{I}$ being the identity.

Then rewrite $\mathcal{N}\boldsymbol{u}_t$ as $\mathcal{N}\boldsymbol{u}_t = (\mathcal{N}\boldsymbol{u})_t - \mathcal{N}_t\boldsymbol{u}$. Notice that $\mathcal{N}$ in equation (2.16) only depends on η and h, but is independent of $\boldsymbol{u}$. Also, h is a function of space and is independent of t. This means $\mathcal{N}_t$ only results in η_t, which can be replaced by $-\nabla\cdot(\eta\boldsymbol{u})$ from equation (2.11). By introducing $U = \mathcal{N}\boldsymbol{u}$, the DSWE in constraint form can be obtained.

DSWE in constraint form in dimension $D = 2$.

$$\eta_t + \nabla\cdot(\eta\boldsymbol{u}) = 0, \tag{2.17}$$

$$\boldsymbol{U}_t + \eta\nabla\zeta + \nabla(\eta\boldsymbol{u}\boldsymbol{u}) = \boldsymbol{F}(\eta,\boldsymbol{u},h), \tag{2.18}$$

$$\boldsymbol{U} - \mathcal{N}\boldsymbol{u} = 0, \tag{2.19}$$

where the forcing $\boldsymbol{F}(\eta, \boldsymbol{u}, h)$ is a nonlinear function given by

$$\begin{aligned}\boldsymbol{F}(\eta, \boldsymbol{u}, h) \equiv \nabla\Big\{ & \frac{1}{3}\eta^3 \left(\boldsymbol{u} \cdot \nabla (\nabla \cdot \boldsymbol{u}) - (\nabla \cdot \boldsymbol{u})^2\right) + \frac{1}{2}\eta^2 (\boldsymbol{u} \cdot \nabla)^2 h \\ & + \eta^2 \nabla \cdot (\eta \boldsymbol{u}) \nabla \cdot \boldsymbol{u} + \eta \nabla \cdot (\eta \boldsymbol{u}) \nabla h \cdot \boldsymbol{u} \Big\} \\ -\nabla h \Big\{ & \frac{1}{2}\eta^2 \left(\boldsymbol{u} \cdot \nabla (\nabla \cdot \boldsymbol{u}) - (\nabla \cdot \boldsymbol{u})^2\right) + \eta (\boldsymbol{u} \cdot \nabla)^2 h \\ & - \eta \nabla \cdot (\eta \boldsymbol{u}) (\nabla \cdot \boldsymbol{u}) + \nabla \cdot (\eta \boldsymbol{u}) (\nabla h \cdot \boldsymbol{u}) \Big\}. \end{aligned} \tag{2.20}$$

In equation (2.18), the term $\boldsymbol{uu}$ is a tensor, where $\boldsymbol{T} = \nabla \cdot (\eta \boldsymbol{uu}) = (T_1, T_2)^T$ with $T_1 = (\eta uu)_x + (\eta uv)_y$ and $T_2 = (\eta uv)_x + (\eta vv)_y$. Here $\boldsymbol{u} = (u, v)$ and $\nabla = (\partial_x, \partial_y)$.

2.3 Simplified Models and Other Related Models of DSWE

For long waves, besides the small long wave parameter β, there are two additional parameters that can be introduced to characterize the shallow water systems:

$$\epsilon \equiv \frac{a}{\bar{h}}, \qquad\qquad \kappa \equiv \frac{h'}{\zeta'}, \tag{2.21}$$

where $h' = O(|\nabla h|)$ and $\zeta' = O(|\nabla \zeta|)$. Here ϵ represents the ratio of the wave amplitude a to the mean water depth $\bar{h}$. The parameter κ measures the variation of depth h regarding to that of the wave height ζ and is assumed to be $O(1)$. As mentioned in Section 2.1, when $\epsilon = O(1)$, the DSWE can be considered as a *strongly nonlinear model*. On the other hand, when $\epsilon \ll 1$, the system can be reduced to a *weakly nonlinear model*. Therefore, equations (2.17)–(2.19) are the strongly nonlinear dispersive shallow water system. Under different approximations of ϵ and κ, (2.17)–(2.19) can be reduced to several simpler systems, which are summarized here.

Case 1: $\kappa \ll 1$, slowly varying bathymetry.

When $|\nabla h| \ll |\nabla \zeta|$ is assumed, the terms involving ∇h in equation (2.18) is negligible so that one can obtain a simplified model from (2.17)–(2.19):

$$\eta_t + \nabla \cdot (\eta \boldsymbol{u}) = 0 \tag{2.22}$$

$$\boldsymbol{U}_t + \eta \nabla \zeta + \nabla(\eta \boldsymbol{u}\boldsymbol{u}) = \nabla \left\{ \frac{1}{3}\eta^3 \left(\boldsymbol{u} \cdot \nabla (\nabla \cdot \boldsymbol{u}) - (\nabla \cdot \boldsymbol{u})^2 \right) + \eta^2 \nabla \cdot (\eta \boldsymbol{u}) \nabla \cdot \boldsymbol{u} \right\}, \tag{2.23}$$

with the PDE constraint:

$$\boldsymbol{U} = \eta \boldsymbol{u} - \nabla \left(\frac{1}{3}\eta^3 \nabla \cdot \boldsymbol{u} \right). \tag{2.24}$$

Even though the variable h no longer appears explicitly in equations (2.22)–(2.24), the bathymetry information is contained in the definition of η as $\eta = h + \zeta$. When the bottom is flat, $\eta = 1 + \zeta$ (in dimensionless form).

Case 2: $\epsilon \ll 1$, weakly nonlinear system.

If the wave amplitude is small compared with the water depth, i.e., $\epsilon \ll 1$, we neglect nonlinear dispersive terms that appear in the definition of $\boldsymbol{U}$ (equation (2.24)) and on the right-hand side of equation (2.23). Then the weakly nonlinear system can be obtained as:

$$\eta_t + \nabla \cdot (\eta \boldsymbol{u}) = 0, \tag{2.25}$$

$$\boldsymbol{U}_t + \eta \nabla \zeta + \nabla(\eta \boldsymbol{u}\boldsymbol{u}) = 0, \tag{2.26}$$

while the PDE constraint becomes

$$\boldsymbol{U} = \eta \boldsymbol{u} - \nabla \left(\frac{1}{3} h^3 \nabla \cdot \boldsymbol{u} \right). \tag{2.27}$$

Notice that the term $\nabla \left(\frac{1}{3} h^3 \nabla \cdot \boldsymbol{u} \right)$ in equation (2.27) still gives equations (2.25)–(2.26) the dispersive effects.

Case 3: The MM69 system (1D).

For comparison, we consider a one-dimensional weakly nonlinear model originally introduced by Madsen and Mei (1969). The Madsen Mei model is referred to as MM69 and will be used for numerical tests.

The MM69 system [11] derived in one dimension under the weakly nonlinear assumption ($\epsilon \ll 1$) is given by

$$\eta_t + (\eta v)_x - \frac{h^3}{6} v_{xxx} = Av + Bv_x + \frac{3}{2} h^2 h_x v_{xx}, \tag{2.28}$$

$$V_t + v v_x + \zeta_x = 0, \tag{2.29}$$

with a PDE constraint

$$V = \left(1 - \left(h_x^2 + h h_{xx}\right)\right) v - 2 h h_x v_x - \frac{h^2}{2} v_{xx}. \tag{2.30}$$

Here the coefficients A, B depend on the a varying bathymetry h, :

$$A = h_x^3 + 3 h h_x h_{xx} + \frac{1}{2} h^2 h_{xxx},$$

$$B = 3 h h_x^2 + \frac{3}{2} h^2 h_{xx}$$

Notice that there is a notable difference between the MM69 and DSWE (and all simplified cases). The DSWE (2.17)–(2.19) (and simplified models) are written in terms of the depth averaged velocity $\boldsymbol{u}$ defined in equation (2.5). Meanwhile, the MM69 system (2.28)–(2.30) is written for the horizontal velocity at the bottom (i.e., $z = -h$). Under the shallow water assumption, the bottom velocity v and depth averaged velocity $\boldsymbol{u}$ become asymptotically equivalent for long wave (as $\beta \to 0$).

Case 4: $h_x \ll 1$ MM69 with slowly varying bathymetry (1D)

Again, under the assumption of slowly varying bathymetry ($\kappa \ll 1$), the MM69 system [11] is simplified to

$$\eta_t + (\eta v)_x - \frac{h^3}{6} v_{xxx} = 0, \tag{2.31}$$

$$V_t + v v_x + \zeta_x = 0, \tag{2.32}$$

with constraint

$$V = v - \frac{h^2}{2} v_{xx}. \tag{2.33}$$

Similarly, the slowly varying bathymetry MM69 system retains the water depth h effect through the relation $\eta = h + \zeta$.

Here for convenience, different models, the underlying assumptions of different models are summarized in Table 2.1.

Table 2.1 Different Models with Their Equations and Assumptions

Equations	Strength of Nonlinearity	Bathymetry Length Scale	Asymptotics
Shallow Water Equations (1D) (A.19)–(A.20)	Strongly Nonlinear System $\epsilon = O(1)$	Flat Bathymetry $h =$ constant	Depth averaged velocity $\bar{\boldsymbol{u}}$ truncated at $O(1)$
Dispersive Shallow Water Equations (DSWE) (2.11)–(2.12)		$\kappa = O(1)$	Depth averaged velocity $\bar{\boldsymbol{u}}$ truncated at $O(\beta^2)$
Reduced DSWE (2.22)–(2.24)		Slowly Varying Bathymetry $\kappa \ll 1$	
Dispersive Weakly Nonlinear Equations (2.25)–(2.27)	Weakly Nonlinear System $\epsilon \ll 1$	Slowly Varying Bathymetry $\kappa \ll 1$	Depth average velocity $\bar{\boldsymbol{u}}$ at $O(\beta^2)$
MM69 System (1D) (2.28)–(2.30)		$\kappa = O(1)$	Horizontal bottom velocity u at $O(\beta^2)$
MM69 System Reduced (1D) (2.31)–(2.33)		Slow Varying Bathymetry $\kappa \ll 1$	

CHAPTER 3

TIME-STEPPING SCHEMES

In this chapter, we present two strategies for numerically time discretizing the DSWE in constraint form. Throughout this chapter, we will leave space continuous and eventually discretize the spatial domain with a spectral method in subsequent chapters.

The primary difficulty involves treating the (nonlinear, time-dependent) operator $\mathcal{N}$ in equations like (2.19). Our goal is to devise time stepping schemes that avoid having to build and/or invert the operator $\mathcal{N}$. In other words, we seek to devise *matrix free* methods, that is, we are willing to perform matrix vector products like $\mathcal{N}(u)$, however, we never construct or invert $\mathcal{N}$ directly.

We provide two time-stepping approaches in this dissertation. The first approach, in Section 3.1, is a standard Runge-Kutta method for the differential equations with a linear solver for the constraint. Here the key point is to find an appropriate preconditioner for the system $\mathcal{N}\boldsymbol{u} = \boldsymbol{U}$ (as $\mathcal{N}\boldsymbol{u} = \boldsymbol{U}$ must be solved at each stage of the Runge-Kutta scheme). In the second time-stepping approach (Section 3.2), a semi-implicit time-discretization is introduced. The main concept is the splitting of the operator $\mathcal{N}$ into an implicit term $\mathcal{A}$, and an explicit term $\mathcal{B}$, thereby resulting in an ImEx scheme where $\mathcal{N} = \mathcal{A} + \mathcal{B}$. The challenge here is that not every choice of $\mathcal{A}$ and $\mathcal{B}$ will result in a stable scheme. We provide theory on how to choose $\mathcal{A}$ and $\mathcal{B}$ so that the resulting scheme is (zero) stable. Finally, we also introduce the One-Leg method corresponding to the semi-implicit ImEx scheme.

3.1 Runge-Kutta Time Discretizations with Preconditioner

In this section, we time discretize the DSWE (such as equations (2.17)–(2.18)) with an explicit Runge-Kutta (RK) method. Here in this dissertation, we will focus on the

well-known fourth order RK4 scheme with Butcher tableau:

$$\begin{array}{c|c} \boldsymbol{c} & \boldsymbol{A} \\ \hline & \boldsymbol{b}^T \end{array} = \begin{array}{c|cccc} 0 & & & & \\ 1/2 & 1/2 & & & \\ 1/2 & 0 & 1/2 & & \\ 1 & 0 & 0 & 1 & \\ \hline & 1/6 & 1/3 & 1/3 & 1/6 \end{array} \tag{3.1}$$

Since the matrix $\boldsymbol{A}$ is lower triangular with zeros along the diagonal, the RK scheme defined by (3.1) is explicit. If we denote $\boldsymbol{w} = (\eta, \boldsymbol{U})^T$, then equations (2.17)–(2.19) can be recast into

$$\boldsymbol{w}_t = \boldsymbol{f}(\boldsymbol{w}, \boldsymbol{u}, t), \tag{3.2}$$

$$g(\boldsymbol{w}, \boldsymbol{u}) = 0, \tag{3.3}$$

where $\boldsymbol{f} = (-\nabla \cdot (\eta \boldsymbol{u}),\ -\eta \nabla \zeta - \nabla(\eta \boldsymbol{u}\boldsymbol{u}) + \boldsymbol{F})^T$ and $\boldsymbol{g} = \boldsymbol{U} - \mathcal{N}\boldsymbol{u}$. Here $\boldsymbol{F}$ is defined in equation (2.20). When applied to equation (3.2), one step of the 4th order RK scheme (3.1) can be written as:

$$\boldsymbol{w}^{(1)} = \boldsymbol{w}_n, \tag{3.4}$$

$$\boldsymbol{w}^{(2)} = \boldsymbol{w}_n + \frac{1}{2}\Delta t\ \boldsymbol{f}\left(\boldsymbol{w}^{(1)}, \boldsymbol{u}^{(1)}, t_n\right), \tag{3.5}$$

$$\boldsymbol{w}^{(3)} = \boldsymbol{w}_n + \frac{1}{2}\Delta t\ \boldsymbol{f}\left(\boldsymbol{w}^{(2)}, \boldsymbol{u}^{(2)}, t_n + \frac{1}{2}\Delta t\right), \tag{3.6}$$

$$\boldsymbol{w}^{(4)} = \boldsymbol{w}_n + \Delta t\ \boldsymbol{f}\left(\boldsymbol{w}^{(3)}, \boldsymbol{u}^{(3)}, t_n + \frac{1}{2}\Delta t\right), \tag{3.7}$$

with the final update:

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \frac{1}{6}\Delta t\Bigg[\boldsymbol{f}\left(\boldsymbol{w}^{(1)}, \boldsymbol{u}^{(1)}, t_n\right) + 2\boldsymbol{f}\left(\boldsymbol{w}^{(2)}, \boldsymbol{u}^{(2)}, t_n + \frac{1}{2}\Delta t\right) + 2\boldsymbol{f}\left(\boldsymbol{w}^{(3)}, \boldsymbol{u}^{(3)}, t_n + \frac{1}{2}\Delta t\right) + \boldsymbol{f}\left(\boldsymbol{w}^{(4)}, \boldsymbol{u}^{(4)}, t_n + \Delta t\right)\Bigg], \tag{3.8}$$

where $\boldsymbol{w}_n \approx \boldsymbol{w}(t_n)$ is the approximation of $\boldsymbol{w}_n$ at t_n, and $t_n = n\Delta t$ when using a fixed time step Δt. Here $\boldsymbol{w}^{(j)}$ are the intermediate stage values (where for brevity we have suppressed the subscript n). The algebraic equation (3.8) can be solved to get the solutions of $\boldsymbol{u}$ with knowledge of the values $(\boldsymbol{U}, \eta)$, (i.e., $\boldsymbol{u} = \mathcal{N}^{-1}\boldsymbol{U}$) where

$$g(\boldsymbol{w}^{(j)}, \boldsymbol{u}^{(j)}) = 0, \qquad \text{for } j = 1, 2, 3, 4 \tag{3.9}$$

is the discretization of the constraint and defines the variable $\boldsymbol{u}^{(j)}$ in terms of $\boldsymbol{w}^{(j)}$.

Inverting the operator $\mathcal{N}$ directly is not computationally attractive because $\mathcal{N}$ depends on η and changes at every point in time. The operator $\mathcal{N}$ is symmetric and positive definite. Hence, the linear system $\mathcal{N}(\boldsymbol{u}) = \boldsymbol{U}$ can be solved using a gradient descent algorithm or conjugate gradient (CG) algorithm [3]. Here a linear operator $\mathcal{A}$ can be introduced into the system as a preconditioner to condition and speed up the algorithm. Then, the preconditioned gradient descent method can solve the constraint equation iteratively (for $n \geq 0$) with an initialized $\boldsymbol{u}_0$:

$$\text{Do:} \qquad \mathcal{A}\boldsymbol{u}_{n+1} = \mathcal{A}\boldsymbol{u}_n + \Delta s(\boldsymbol{U} - \mathcal{N}\boldsymbol{u}_n), \tag{3.10}$$

$$\text{until:} \qquad \frac{\|\boldsymbol{u}_{n+1} - \boldsymbol{u}_n\|}{\|\boldsymbol{u}_n\|} < \delta. \tag{3.11}$$

$\Delta s > 0$ is a pseudo-time. Here $\|\boldsymbol{u}\| = \int_\Omega u^2(x)dx$ and the small parameter δ in the termination criteria (3.11) appears as the tolerance to stop the solver when the iterative scheme is applied repeatedly.

We now provide a few details on the convergence of equation (3.10). For linear operators $\mathcal{A}$ and $\mathcal{N}$, we can introduce the generalized eigenvalues and eigenvectors μ and $\boldsymbol{v}$ where $\mathcal{A}\boldsymbol{v} = \mu\mathcal{N}\boldsymbol{v}$. If $\mathcal{A}$ is positive definite and $\mathcal{N}$ is positive semi-definite, the eigenvalues μ are real, non-negative and bounded, i.e., $0 \leq \mu < \infty$. Let $\boldsymbol{e}_n = \boldsymbol{u}_n - \boldsymbol{u}$ be the error in the gradient descent (3.10) where $\boldsymbol{u}$ is the exact solution to $\boldsymbol{U} = \mathcal{N}\boldsymbol{u}$. If we let $\boldsymbol{e}_n = e_n\boldsymbol{v}$ where $\boldsymbol{v}$ is a generalized eigenvector of $\mathcal{A}$ and $\mathcal{N}$ (with eigenvalue μ). Here $R \equiv \max\{|1 - \Delta s\mu| : \mu \text{ are eigenvalues of } \mathcal{A}^{-1}\mathcal{N}\}$ [24]. Then, e_n satisfies the scalar equation $e_{n+1} = re_n$, where $r = (1 - \Delta s\mu)$. Hence $e_n = r^n e_0$. It follows that provided $-1 < r < 1$, each mode $\boldsymbol{e}_n = e_n\boldsymbol{v}$ converges to zero (i.e., (3.10) is a contraction mapping). In other words, we obtain convergence so long as $\Delta s < 2/\mu$ is satisfied for every eigenvalue μ of $\mathcal{A}^{-1}\mathcal{N}$ (this will then ensure $-1 < r < 1$).

It is also worth mentioning that the stopping criteria (3.11) is in fact achieved. Assuming $R < 1$, for large n, (3.10) becomes:

$$\frac{\|\boldsymbol{u}_{n+1} - \boldsymbol{u}_n\|}{\|\boldsymbol{u}_n\|} \approx \frac{\|\boldsymbol{e}_{n+1} - \boldsymbol{e}_n\|}{\|\boldsymbol{u}_n\|} \sim R^n. \tag{3.12}$$

The termination criteria (3.11) then becomes (approximately) $R^n < \delta$, which is achieved when $n \sim \ln(\delta)/\ln(R)$.

Now, we turn our attention to discuss the choice of $\mathcal{A}$ in (3.10) that optimizes the convergence rate R. The conditioning number of a positive definite matrix $\boldsymbol{A}$ is $\kappa(\boldsymbol{A}) = \mu_{\max}/\mu_{\min}$, where $\mu_{\max}$ and $\mu_{\min}$ are the largest and smallest eigenvalues of $\boldsymbol{A}$. For an optimal choice of $\Delta s = 2(\mu_{\min}+\mu_{\max})^{-1}$, where $\mu_{\min}$ and $\mu_{\max}$ are the minimum and maximum eigenvalues of $\mathcal{A}^{-1}\mathcal{N}$, the convergence rate R can be written in terms of the conditioning number $\kappa(\mathcal{A}^{-1}\mathcal{N})$. Our choice of preconditioner will need to ensure

a bound on the conditioning number $\kappa(\mathcal{A}^{-1}\mathcal{N})$. Ideally, the conditioning number $\kappa(\mathcal{A}^{-1}\mathcal{N})$ should be independent of the (subsequent) mesh discretization. This will then ensure that the iterative solver remains efficient, even at fine mesh resolutions ($\Delta x \to 0$). Since $\mathcal{N}$ is symmetric and positive definite ($\mathcal{N} \succ 0$), the convergence rate of the preconditioned gradient method is $R = (\kappa - 1)/(\kappa + 1) \sim e^{-1/\kappa}$ which assumes an optimally chosen step size. Therefore, the estimated number of iterations required to achieve the tolerance δ is $O(\kappa)$. Note that without the preconditioner, the conditioning number $\kappa(\mathcal{N})$ will slow down the iterative scheme or eventually fail with small mesh size.

Note: RK4 is a multistage time-stepping method. At each intermediate stage, the preconditioning gradient descent scheme should be applied to solve for the values of $\boldsymbol{u}$ for next stage.

Here we summarize the convergence rates, and necessary conditions for the fixed point iteration (FPI) (3.10) and conjugate gradient methods to solve systems of the form $\mathcal{N}\boldsymbol{a} = \boldsymbol{b}$. In Table 3.1, $\kappa = \kappa(\mathcal{A}^{-1}\mathcal{N})$; $\sigma(\mathcal{A}^{-1}\mathcal{N})$ is the spectrum of $\mathcal{A}^{-1}\mathcal{N}$; and e_i is the error: $e_i = \|\boldsymbol{a}_i - \boldsymbol{a}\|_{\mathcal{N}}$ for CG and $e_i = \|\boldsymbol{a}_i - \boldsymbol{a}^*\|_2$ for FPI.

Table 3.1 Summary of Convergence Rates

	Fixed point iteration (FPI)	Conjugate gradient (CG)
Formula	Initialize: $\boldsymbol{a}_0$. For $i \geq 0$ $\mathcal{A}\boldsymbol{a}_{i+1} = \mathcal{A}\boldsymbol{a}_i - (\mathcal{N}\boldsymbol{a}_i - \boldsymbol{b})$	See [24, Chapter 38] for algorithm
Conditions for use	$R \equiv \max\{\|1-\mu\| : \mu \in \sigma(\mathcal{A}^{-1}\mathcal{N})\}$, then $R < 1$ is required. If $\mathcal{N} \succ 0, \mathcal{A} \succ 0$, $R < 1$ is $\mu_{\max}(\mathcal{A}^{-1}\mathcal{N}) < 2$.	$\mathcal{G} \succ 0, \mathcal{A} \succ 0$, No condition on $\mu_{\max}(\mathcal{A}^{-1}\mathcal{N})$
Convergence rate bound $e_i \leq CR^i e_0$	$R = \max\{\|1-\mu\| : \mu \in \sigma(\mathcal{A}^{-1}\mathcal{N})\}$, $C = 1$ Optimal case: If $\mathcal{A} \succ 0$, $\mathcal{N} \succ 0$, and $\|1-\mu_{\max}\| = \|1-\mu_{\min}\|$, then $R = \dfrac{\kappa-1}{\kappa+1} \sim e^{-1/\kappa}$	$C = 2$ $R = \dfrac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \sim e^{-1/\sqrt{\kappa}}$
Estimate number of iterations required	$\sim O(\kappa)$ (In optimal case)	$\sim O(\sqrt{\kappa})$

We did not pursue implementing conjugate gradient, however we expect that conjugate gradient method will further decrease the number of iterations (as it converges with a rate of $O(\sqrt{\kappa})$ as opposed to $O(\kappa)$).

3.2 ImEx Methods

In this section, we introduce a semi-implicit (ImEx) method. The idea is to split the operator $\mathcal{N}$ into an implicit and explicit part to avoid a fully implicit treatment. The splitting can be done as $\mathcal{N} = \mathcal{A}+\mathcal{B}$ where $\mathcal{A}$ is the implicit term (Im) and $\mathcal{B} = \mathcal{N}-\mathcal{A}$ is the explicit term (Ex). Since $\mathcal{A}$ is treated implicitly, $\mathcal{A}$ is picked to be simpler for fast computational time.

In this dissertation, a second order backward differential formula is used along with the ImEx method. One thing that needs to be checked of this scheme is zero stability, which is a key requirement for convergence. Normally zero stability is easy to verify; however, here it's more difficult due to the ImEx structure. Since the DSWE are restructured into the DAE form, then the zero stability depends on both the ImEx scheme and the operator $\mathcal{N}$ (choice of $\mathcal{A}$ and $\mathcal{B}$). Similar to [18], an ImEX scheme can be established with an appropriate set of time-stepping coefficients to ensure stability. Under the notation of $\boldsymbol{w} = (\eta, \boldsymbol{U})^T$, the general ImEx multistep method can be applied to equations (3.2)–(3.3) as

$$\frac{1}{\Delta t}\sum_{j=0}^{s} a_j \boldsymbol{w}^{n+j} = \sum_{j=0}^{s} b_j \boldsymbol{f}(\boldsymbol{w}^{n+j}, \boldsymbol{u}^{n+j}, t^{n+j}), \tag{3.13}$$

$$\sum_{j=0}^{s} c_j \mathcal{A}\boldsymbol{u}^{n+j} + b_j \mathcal{B}_{n+j}\boldsymbol{u}^{n+j} = \sum_{j=0}^{s} c_j \boldsymbol{U}^{n+j}, \tag{3.14}$$

where $\mathcal{B}_n = \mathcal{N}_n - \mathcal{A}$. The *ImEx coefficients* (a_j, b_j, c_j), $0 \le j \le s$, are chosen with $c_s \neq 0$, $b_s = 0$ so that equation (3.14) is implicit in $\mathcal{A}$ and explicit in $\mathcal{B}$. Here the coefficients (a_j, b_j, c_j) should be picked to satisfy order conditions to ensure that

equations (3.13)–(3.14) define an second order time-stepping scheme (second order backward differential scheme in this dissertation) [19]. By using a Taylor expansion, the time-stepping coefficients can be easily chosen (similar to backward differential formula 2) to get the numerical scheme:

$$\frac{1}{\Delta t}\left(\frac{3}{2}\boldsymbol{w}_{n+1} - 2\boldsymbol{w}_n + \frac{1}{2}\boldsymbol{w}_{n-1}\right) = 2\boldsymbol{f}(\eta_n, \boldsymbol{u}_n, t_n) - \boldsymbol{f}(\eta_{n-1}, \boldsymbol{u}_{n-1}, t_{n-1}), \tag{3.15}$$

where the constraint can be solved as

$$\mathcal{A}\boldsymbol{u}_{n+1} + (2\mathcal{B}_n\boldsymbol{u}_n - \mathcal{B}_{n-1}\boldsymbol{u}_{n-1}) = \boldsymbol{U}_{n+1}. \tag{3.16}$$

Equations (3.15)–(3.16) are the full multistep scheme where the operator $\mathcal{B}$ and $\boldsymbol{u}$ are treated are treated at the same time values when applying the second order scheme. There is also a second order One-Leg method [7], which is related to the multistep method, where $\mathcal{B}$ and $\boldsymbol{u}$ are not treated at the same time values. Since the operator $\mathcal{B}$ contains $\eta(\boldsymbol{x}, t)$ in it, then η and $\boldsymbol{u}$ can be treat individually to achieve the second order scheme. At this time, we remark that one could also time discretize the DSWE directly without first recasting them into a DAE form. Specifically, equation (2.12) can be rewritten as:

$$\widehat{\mathcal{N}}(\boldsymbol{u}_t) = \boldsymbol{F}(\eta, \boldsymbol{u}, t). \tag{3.17}$$

The operator $\widehat{\mathcal{N}}$ now has the form

$$\widehat{\mathcal{N}} = I - \frac{1}{\eta}\nabla\left\{\left(\frac{\eta^3}{3}\nabla\cdot\right) - \left(\frac{\eta^2}{2}\nabla h\cdot\right)\right\} + \nabla h\left\{\left(\frac{\eta}{2}\nabla\cdot\right) + (\nabla h\cdot)\right\}, \tag{3.18}$$

where $\boldsymbol{F}(\eta, \boldsymbol{u}, t)$ contains the rest of terms in equation (2.12). The operator $\widehat{\mathcal{N}}$ here can still be split into two parts — $\widehat{\mathcal{N}} = \widehat{\mathcal{A}} + \widehat{\mathcal{B}}$. Compare to ImEx operators and splitting matrices, $\mathcal{N} = \eta\widehat{\mathcal{N}}$, $\mathcal{A} = \eta\widehat{\mathcal{A}}$ and $\mathcal{B} = \eta\widehat{\mathcal{B}}$. Again, like the original ImEx method which discussed before, $\widehat{\mathcal{A}}$ is treated implicitly and $\widehat{\mathcal{B}}$ is treated explicitly. By using the Taylor's expansion, a second order One-Leg scheme to solve equations (2.11)–(2.12) can be obtained as

$$\frac{\eta^{n+1} - \eta^n}{\Delta t} = -\nabla \cdot \left[\eta^{n+\frac{1}{2}} \boldsymbol{u}^{n+\frac{1}{2}}\right], \tag{3.19}$$

$$\widehat{\mathcal{A}} \frac{\boldsymbol{u}^{n+1} - \boldsymbol{u}^n}{\Delta t} = \widehat{\mathcal{B}}_{\eta^{n+\frac{1}{2}}} \frac{2\boldsymbol{u}^n - 3\boldsymbol{u}^{n-1} + \boldsymbol{u}^{n-2}}{\Delta t} + \boldsymbol{F}\left(\eta^{n+\frac{1}{2}}, \boldsymbol{u}^{n+\frac{1}{2}}, t + \frac{\Delta t}{2}\right). \tag{3.20}$$

Here $\eta^{n+\frac{1}{2}}$ and $\boldsymbol{u}^{n+\frac{1}{2}}$ are

$$\boldsymbol{u}^{n+\frac{1}{2}} = \frac{3}{2}\boldsymbol{u}^n - \frac{1}{2}\boldsymbol{u}^{n-1}, \tag{3.21}$$

$$\eta^{n+\frac{1}{2}} = \frac{3}{2}\eta^n - \frac{1}{2}\eta^{n-1}. \tag{3.22}$$

Since this method is no longer a multistep ImEx method (it is a one-leg version of an ImEx scheme), yet still has explicit parts and implicit parts, we refer to (3.19)–(3.22) as a second order semi-implicit method.

Choosing the ImEx (or semi-implicit) splitting of the operator $\mathcal{N}$ (or $\hat{\mathcal{N}}$) can be tricky. The ImEx (or semi-implicit) method here treats parts of the operator $\mathcal{N}$ (or $\hat{\mathcal{N}}$) explicitly — and these explicit terms which consist of $\mathcal{B}$ (or $\hat{\mathcal{B}}$) may be stiff. Specifically, for our preferred choice of splitting, some terms in the matrix $\mathcal{B}$ (or $\hat{\mathcal{B}}$) will contain higher order derivatives, and these terms can cause the scheme to become unstable. For the ImEx (or semi-implicit schemes) zero stability becomes a crucial property required to ensure that the splitting of the operator $\mathcal{N}$ (or $\hat{\mathcal{N}}$) is stable. Zero stability will be ensured provided the implicit term $\mathcal{A}$ stabilizes the explicit term $\mathcal{B}$.

We now state conditions for zero stability in a simplified setting where the operator $\mathcal{N}$ is assumed allowed to vary in space, but is constant in time (which is generally not the case as η depends on both space and time). With this assumption, we seek splittings $\mathcal{A}$ and $\mathcal{B}$ where $\mathcal{A}$ is constant coefficient, $\mathcal{B}$ is allowed to be variable coefficient and both $\mathcal{A}$ and $\mathcal{B}$ do not depend on time. Let the generalized eigenvalues of $\mathcal{A}$ and $\mathcal{B}$ be:

$$\mathcal{A}\boldsymbol{u} = \mu\mathcal{B}\boldsymbol{u}. \tag{3.23}$$

In equation (3.23) both and $\mathcal{A}$ and $\mathcal{B}$ are understood to have periodic boundary conditions, and we assume that $\mathcal{A}$ is invertible so that μ is never an eigenvalue.

To formulate the zero-stability of the ImEx splitting will examine solutions to equation (3.20) with $F = 0$. Then $\boldsymbol{u}^n$ can be expressed as $\boldsymbol{u}^n = z^n\boldsymbol{u}_0$ where $z \in \mathbb{C}$. With this substitution in (3.20), we have the recurrence:

$$\hat{\mathcal{A}}(z^3u_0 - z^2u_0) = \hat{\mathcal{B}}(2z^2 - 3z + 1)u_0, \tag{3.24}$$

$$\phi(z;\mu)\hat{\mathcal{A}}\boldsymbol{u}_0 = 0, \tag{3.25}$$

where here we have used the introduction of μ to be a generalized eigenvalue of $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ and $\phi(z;\mu)$ is the polynomial given by:

$$\phi(z;\mu) = (z-1)(z^2 - 2z\mu + \mu). \tag{3.26}$$

The modes $\boldsymbol{u}_0$ are stable if z satisfies the root condition:

Definition 3.1. *(Root condition) A polynomial $p(z)$ satisfies the root condition if each simple each root is $|z| \leq 1$, and non-simple roots satisfy $|z| < 1$.*

For stability, we need that every root of $\phi(z;\mu)$ satisfy the root condition. To do so we introduce a geometric stability diagram.

Definition 3.2. *(Zero-stability region) The values of μ that ensure the root condition:* $D := \{\mu \in C : \phi(z;\mu) haseachroot |z| < 1\}$

We then have the following condition for zero-stability. The roots to equation (3.26), are stable if for every generalized eigenvector μ (in equation (3.23)) lies in $\mu \in D$.

To compute the zero-stability diagram, which corresponds to z that lie inside the unit circle, we use the boundary locus method [9] to obtain the region for μ. The boundary locus for equation (3.20) is:

$$\Gamma := \{\mu = z^2/(2z-1) : z = e^{i\theta}, 0 \in [0, 2\pi])\}. \tag{3.27}$$

Figure 3.1 shows the plot of Γ for the second order method.

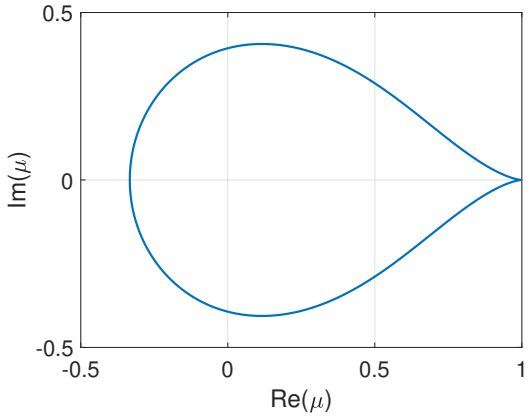


Figure 3.1 Zero stablity region for generalized eigenvalus μ of the second order ImEx (or semi-implicit) method.

Note here, only if the eigenvalues of the ImEx (or semi-implicit) scheme lie inside region, which is shown in Figure 3.1, can the scheme be zero-stable. Ensuring

that the eigenvalues lies inside this stability region places a restriction when splitting the operator $\mathcal{N}$ (or $\hat{\mathcal{N}}$).

3.3 Formulation of $\mathcal{A}$

For the sake of computational efficiency, we seek a constant coefficient symmetric positive definite $\mathcal{A}$ in both the preconditioned gradient descent and the ImEx method. On a periodic domain, the constant coefficient operator $\mathcal{A}$ will be easy to invert using the Fourier transform. Furthermore, in Sections 3.1 and 3.2, we assume that $\mathcal{N}$ is a symmetric positive definite operator. These assumptions will need to be verified.

Notice that the DAE framework ends up producing numerical schemes that are (very close to) direct ImEx discretizations of the original DSWE, hence they are easy to implement. Here we go through several of the simplified DSWE models and discuss our choice of $\mathcal{A}$ for the different $\mathcal{N}$:

Case 1: Slow Varying Bathymetry $\nabla h \ll \nabla \eta$.

Equation (2.24) has the form $\boldsymbol{U} = \mathcal{N}_\eta \boldsymbol{u}$, in which case we take $\mathcal{A}$ to be of the form $\mathcal{A} = \nu I - \alpha \nabla\nabla\cdot$ (which is the linearized operator $\mathcal{N}_\eta$). The two unknown coefficients (ν, α) are chosen under the constraint of zero-stability. The eigenvalue μ, which are obtained as $\mathcal{A}^{-1}\mathcal{N}_\eta \boldsymbol{u} = \mu \boldsymbol{u}$, should lie inside the zero-stability region to guarantee the stability of the corresponding schemes. Both $\mathcal{A}$ and $\mathcal{N}_\eta$ are symmetric and positive definite, hence μ are real, positive and bounded by the Rayleigh quotients:

$$\mu_{\min} = \min_{\boldsymbol{u}\in[H^1(\Omega)]^2} \frac{\langle \boldsymbol{u}, \mathcal{N}_\eta \boldsymbol{u}\rangle}{\langle \boldsymbol{u}, \mathcal{A}\boldsymbol{u}\rangle}\,, \qquad \mu_{\max} = \max_{\boldsymbol{u}\in[H^1(\Omega)]^2} \frac{\langle \boldsymbol{u}, \mathcal{N}_\eta \boldsymbol{u}\rangle}{\langle \boldsymbol{u}, \mathcal{A}\boldsymbol{u}\rangle}\,. \tag{3.28}$$

To ensure the stability, $\mathcal{A}$ can be chosen so that the ratio in equations (3.28) can be bounded as $[\mu_{\min}, \mu\max] \subseteq [0, 1]$ (so that the eigenvalues lie inside the region given by Figure 3.1). A possible choice is $\mathcal{A} = \eta_{\max}\boldsymbol{I} - \frac{1}{3}\eta_{\max}^3 \nabla(\nabla\cdot)$ where $\eta_{\max} = \max_{\boldsymbol{x}} \eta(\boldsymbol{x})$. The efficiency of this matrix $\mathcal{A}$ is shown in Figure 3.2.

Case 2: Bathymetry with constant (non-neglectable) slope.

When ∇h is non-neglectable, the operator $\mathcal{N}$ contains a symmetric, positive definite $\mathcal{N}_\eta$ and a (possibly) indefinite part $\mathcal{K}_{\eta,h}$ where $\mathcal{N} = \mathcal{N}_\eta + \mathcal{K}_{\eta,h}$ and

$$\mathcal{K}_{\eta,h} \equiv -\nabla\left(\tfrac{1}{2}\eta^2\nabla h^T\cdot\right) + \tfrac{1}{2}\eta^2\nabla h(\nabla\cdot) + \eta\nabla h(\nabla h^T\cdot). \tag{3.29}$$

The symmetric property of operator $\mathcal{K}$ can be proved by integrating by parts $\langle\mathcal{K}\boldsymbol{u},\boldsymbol{v}\rangle = \langle\boldsymbol{u},\mathcal{K}\boldsymbol{v}\rangle$. But the positive definite property of $\mathcal{K}$ cannot be guaranteed. To gain insight as to whether $\mathcal{N}$ (or equivalently $\hat{\mathcal{N}}$) is positive definite, we examine $\mathcal{N}$ under the simplified assumption that $\boldsymbol{w} = \nabla h$ is constant and η is constant on an infinitely large domain. Here $\boldsymbol{w} = \nabla h$ is constant means there is a constant sloped bathymetry. Then, under this simplified case, the operator $\mathcal{N}$ is (block) diagonal in Fourier space:

$$\widetilde{\mathcal{N}} = \eta\boldsymbol{I} + \eta\boldsymbol{w}\boldsymbol{w}^T + \tfrac{1}{3}\eta^3\boldsymbol{k}\boldsymbol{k}^T + \tfrac{1}{2}i\eta^2(\boldsymbol{w}\boldsymbol{k}^T - \boldsymbol{k}\boldsymbol{w}^T), \tag{3.30}$$

where $\hat{\mathcal{N}}(\boldsymbol{k})$ is the Fourier transformed operator, and $\boldsymbol{k}$ are the wave numbers. Writing $\boldsymbol{k} = \|\boldsymbol{k}\|(\cos\theta, \sin\theta)$, the eigenvalues of $\hat{\mathcal{N}}(\boldsymbol{k})$ in equation (3.30) can be obtained as

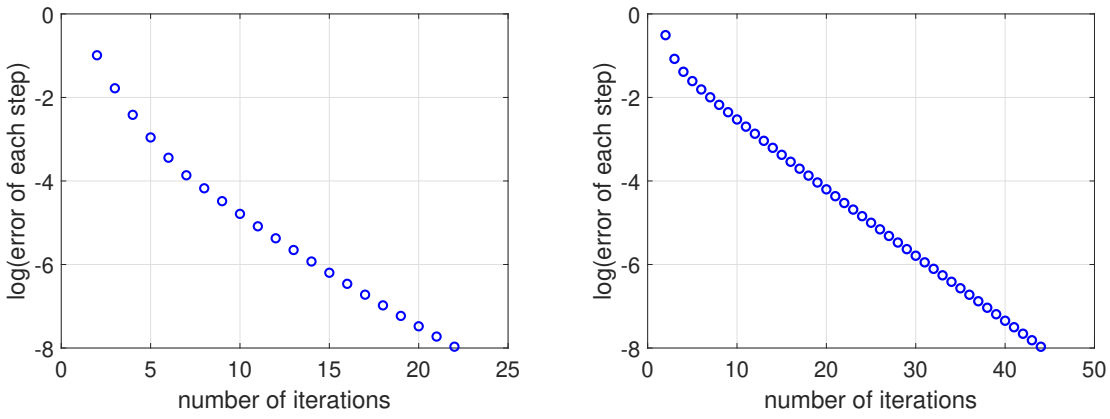
$$\mathrm{eig}(\widetilde{\mathcal{N}}) = \eta + \eta\|\boldsymbol{w}\|^2\lambda_\pm, \tag{3.31}$$

with

$$\lambda_\pm = \tfrac{1}{2}\left(\tfrac{1}{3}\tau^2 + 1\right) \pm \tfrac{1}{2}\left(\left(\tfrac{1}{3}\tau^2 + 1\right)^2 - \tfrac{1}{3}\tau^2\sin^2\theta\right)^{\frac{1}{2}}. \tag{3.32}$$

Here $\tau = \eta\|\boldsymbol{k}\|\|\boldsymbol{w}\|^{-1}$. One has $\lambda_{\pm} > 1$ for all τ and θ; so $\mathcal{N}$ is positive definite. Choosing $\mathcal{A} = (\eta_{\max} + \|\nabla h\|^2_{\max})\boldsymbol{I} - \frac{1}{3}\eta^3_{\max}\nabla(\nabla\cdot)$ can ensure a zero-stable second order scheme.

Figure 3.2 shows the convergence rate of the iterative schemes while solving the constraint equations. Figure 3.2 (a) shows the linear convergence rate of iterative scheme under the slow varying bathymetry assumption (equation (2.24)) and Figure 3.2 (b) represents that under bathymetry with constant non-neglectable gradient (equation (2.19)). The errors were all computed in the ℓ_∞ norm defined on the computational grid (we used an equispaced periodic grid with $N = 256$ gridpoints).



(a) Slow varying bathymetry DSWE (b) DSWE

Figure 3.2 Solving the constraint equation with iterative schemes to get the error of each iteration.

CHAPTER 4

TEST CASES AND RESULTS

In this chapter, we present numerical results of the DSW equations for the propagation of a single solitary wave propagation over bottom topography. After introducing $\boldsymbol{\xi} = x\cos(\theta) + y\sin(\theta) - ct$ with an arbitrary angle θ, the solitary wave solution of the DSWE can be written as

$$\zeta = a\,\mathrm{sech}^2(\gamma\boldsymbol{\xi}). \tag{4.1}$$

When θ is chosen to be 0 or $\pi/2$, the solitary wave is traveling in the x or y direction. Here a is the wave amplitude, c is the wave speed given by $c = \sqrt{1+a}$, and γ is given by $\gamma = \frac{\sqrt{3(c^2-1)}}{2c}$. The velocity vector $\boldsymbol{u}$ can be written as $\boldsymbol{u} = (w\cos\theta, w\sin\theta)$, where w is given by

$$w = \frac{c\zeta}{\zeta + h}. \tag{4.2}$$

First, in Section 4.1, both the RK4 method with preconditioner and the ImEx method are tested for flat bottm. The efficiency of these two methods is examined. In Section 4.2, two different bottom profiles are considered for the DSWE system. After investigating the convergence rates of the two methods, the deformation of a single solitary wave is studied when it's propagating over a constant slope topography.

In this dissertation, we discretize the computational domain with an equispaced grid ($N = 256$) and impose periodic boundary conditions. The grid size is $h = \Delta x = \Delta y$ with grid points $x_i = ih$; $y_j = hj$ and $0 \le i, j \le N-1$. Then define $u_{ij} = u(x_i, y_j)$. To compute spatial derivatives, we use a (standard) Fourier spectral method. The

discrete Fourier transform [23] of a single variable function $u_i = u(x_i)$ is

$$\hat{u}_k = h \sum_{j=1}^{N} e^{-ikx_j} u_j, \qquad k = -\frac{N}{2} + 1, ..., \frac{N}{2}, \tag{4.3}$$

and the inverse discrete Fourier transform is given by:

$$u_j = \frac{1}{2\pi} \sum_{k=-N/2+1}^{N/2} e^{ikx_j} \hat{u}_k, \qquad j = 1, ...N. \tag{4.4}$$

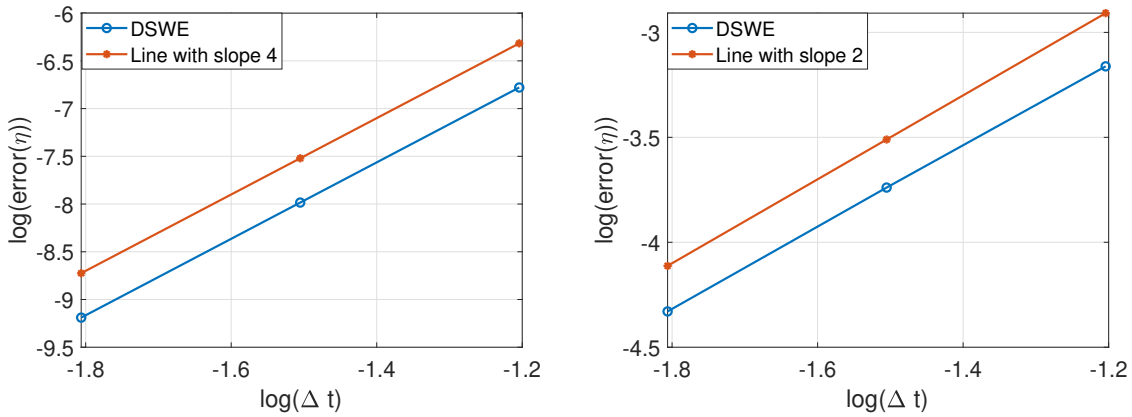
Then, the $w = n^{th}$-order derivative of function u is

- Compute $\hat{u}$.
- Define $\hat{w}_k = (ik)^n \hat{u}_k$, but $\hat{w}_{n/2} = 0$ if n is odd.
- Compute w form $\hat{w}$.

4.1 Flat Bottom Case

For the flat bottom (or $h = 1$ in dimensionless form), all terms with ∇h vanish so that DSWE (2.17)–(2.19) can be simplified to equations (2.22)–(2.24). Here $\eta = 1 + \zeta$. Also, the operator $\mathcal{A}$ used here is $\mathcal{A} = \eta_{\max} \boldsymbol{I} - \frac{1}{3}\eta_{\max}^3 \nabla(\nabla\cdot)$.

For the propagation of a solitary wave, the fourth-order convergence rate of the RK4 method with preconditioner and the second-order convergence rate of the ImEx method are shown in Figure 4.1, as expected.



(a) RK4 method with preconditioner (b) Second-order ImEx method

Figure 4.1 Convergence rate of the numerical schemes for the flat bottom. Here the solitary wave amplitude is $a = 0.5$.

One can see from the Figure 4.1 that the operator $\mathcal{A}$ chosen in Section 3.3 seems to be suitable for the fixed point iteration method as well as the ImEx method. The iteration scheme affects little the convergence rate of the Runge-Kutta time-stepping method. In addition, the ImEx method stays stable with our choice of $\mathcal{A}$.

Another property of these two methods to be tested is their numerical efficiency. For a fair comparison, a second-order Runge-Kutta with preconditioner method should be adopted to compare with the second-order ImEx method. Figure 4.2 shows the computation times of the two methods versus their errors. It can be observed that the ImEx method takes less time than that of the RK2 with preconditioner method for the same order of accuracy. Based on this results, one can conclude that the ImEx method is more efficient.

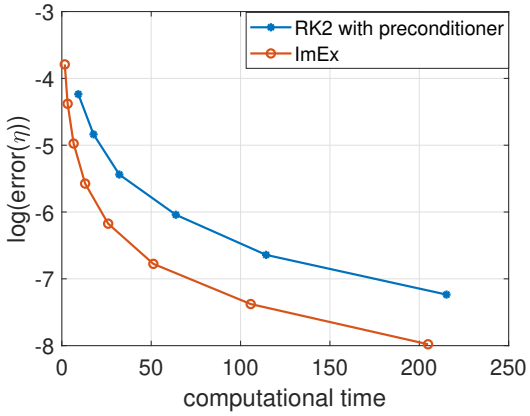


Figure 4.2 Comparison between two second order schemes.

4.2 Variable Bottom Case

As the two numerical methods are validated for the flat bottom, we now consider two different topographies for the DSWE system. Since the pseudo-spectral method is used for spatial discretization, the topographies must be periodic and C^∞. As it's defined in Figure 2.1, the topography is represented by a known function $f(\boldsymbol{x})$ so that the local water depth is defined as $h(\boldsymbol{x}) = 1 - f(\boldsymbol{x})$ (dimensionless form). After the computational domain is chosen to be $\Omega = (-40\pi, 40\pi) \times (-40\pi, 40\pi)$, a Gaussian function is assumed for the first bottom topography (a 'bump' is located at the center of the domain), where the Gaussian function itself is C^∞ and periodic (Figure 4.3(a)):

$$f(\boldsymbol{x}) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{\|\boldsymbol{x}\|^2}{2\sigma^2}}. \tag{4.5}$$

Here $\|\boldsymbol{x}\|$ is the Euclidean norm and $\sigma = \sqrt{3.5}$. This topography is used to test the convergence rate of the schemes in this research.

Then the propagation of a simple solitary wave passing over a constant-slope topography (a 'ramp') is studied numerically to compare the DSWE with the system

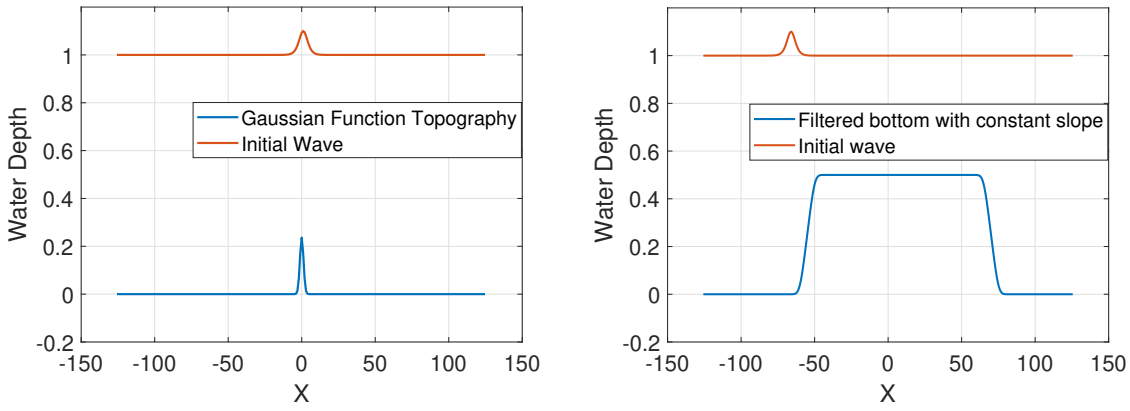
of MM69 (Figure 4.3(b)). The bottom profile $f(\boldsymbol{x})$ is given by

$$f(x) = \begin{cases} \frac{1}{20}(x+60), & -60 \le x \le -50 \\ 0.5, & -50 < x \le 65 \\ -\frac{1}{20}(x-75), & 65 < x \le 75 \\ 0. & \text{elsewhere} \end{cases} \tag{4.6}$$

To enforce the periodic and C^∞ properties, a Gaussian filter defined by

$$\hat{f}(x) = \mathcal{F}^{-1}\left[\mathcal{F}\left[f(x)\right] g(k)\right], \tag{4.7}$$

is applied to the topography function. Here $\mathcal{F}$ and $\mathcal{F}^{-1}$ are Fourier and inverse Fourier transform, $g(k)$ is a Gaussian function. Then, the filtered topography $\hat{f}(x)$ is smoothed out. The initial locations of a solitary wave related to the two topographies are shown in Figure 4.3.



(a) Guassian function topography (b) Constant slope topography

Figure 4.3 Initial solitary waves for the different topographies.

For the variable bottom case, the solitary wave solution given by equation (4.1) is not the exact solution to the DSWE so that a manufacturing method is used to test the convergence rate of the numerical schemes discussed in this dissertation. The idea of the manufacturing method is that, before solving a partial differential equation (PDE) given by

$$u_t = f(u, t), \tag{4.8}$$

for which no exact solution is know, we consider a modified equation for which we can choose any arbitrary function as an exact solution. After choosing arbitrarily u^* as a solution, a forcing term $f^*(x, t)$ is added to the equation (4.8), where the forcing term is given by

$$f^*(x, t) = u_t^* - f(u^*, t). \tag{4.9}$$

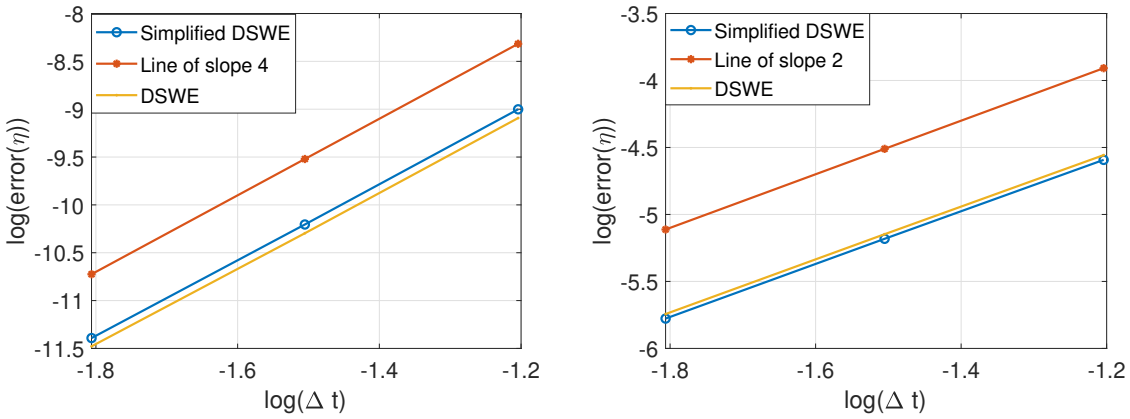
Then, instead of solving equation (4.8), we solve a new PDE

$$u_t = f(u, t) + f^*(x, t), \tag{4.10}$$

whose exact solution is $u^*(x, t)$. Then, the convergence rate can be found by comparing the numerical solution of equation (4.10) with the exact solution u^*.

With the manufacturing method, the time-stepping methods discussed in Chapter 3 can be tested for the variable bottom case. For the Gaussian topography, Figure 4.4 (a) shows the fourth-order convergence rate of the RK4 method with the preconditioner given by equations (3.8), (3.10), and (3.11) while Figure 4.4 (b) represents the second-order convergence rate of the ImEx method given by equations

(3.15)–(3.16). We also provide the results of the simplified DSWE ((2.22)–(2.24)) solved by these two methods.



(a) RK4 method with preconditioner (b) Second-order ImEx method

Figure 4.4 Convergence rates of the numerical schemes.

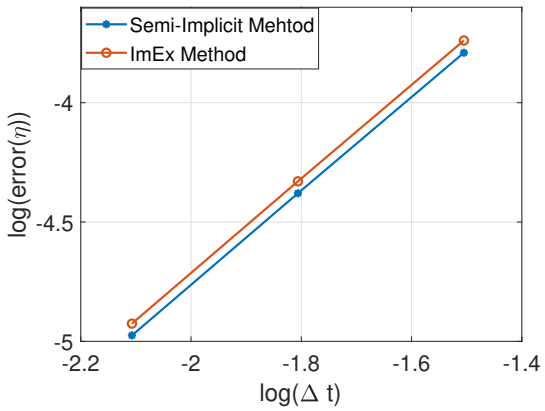


Figure 4.5 Second-order convergence rate of semi-implicit method and ImEx method.

Figure 4.5 shows the second-order convergence rate of the semi-implicit method with equations (3.19)–(3.20). It can be seen that both the ImEx method and the semi-implicit method are second-order accurate, as expected, and their results are comparable.

Using both the RK method with preconditioner and the ImEx method, we next study the propagation of a solitary wave over the constant-slope topography ('ramp'

bottom in Figure 4.3(b)). When the wave is propagating over the ramp, the initial solitary wave is disintegrated into a few solitary waves as shown in Figure 4.6. We also provide the propagation of solitary wave in simplified DSWE system. From Figure 4.7, one can also see that the DSWE system is little affected by the terms with ∇h.

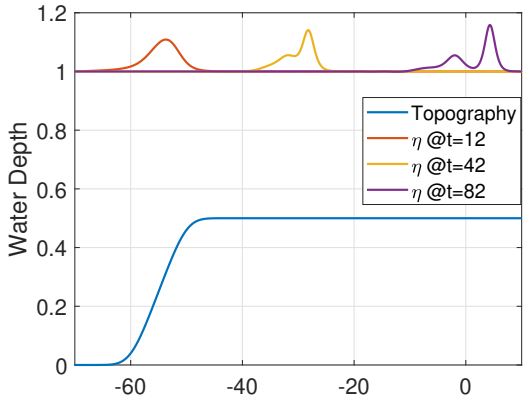


Figure 4.6 A solitary wave propergating over the topography shown in Figure 4.3(b) in DSWE system. The amplitude of the initial wave is chosen to be $a = 0.1$.

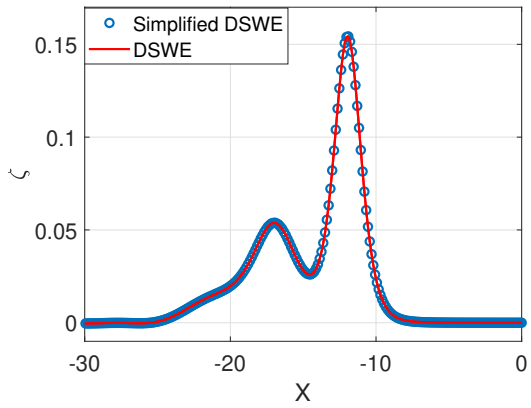


Figure 4.7 Simplified DSWE vs DSWE when wave propagating over the bathymetry shown in Figure 4.3(b) at $t = 62$.

Additional tests are performed for the comparison of the DSWE with the weakly nonlinear model (2.25)–(2.27) and MM69 (2.28)–(2.30). For the two weakly nonlinear models, the wave amplitude is chosen to be 0.1 to satisfy the weakly nonlinear assumption ($\epsilon \ll 1$). As they have different dispersion relations, the numerical

solutions of the two models show some discrepancy in phase, but the overall agreement is reasonable.

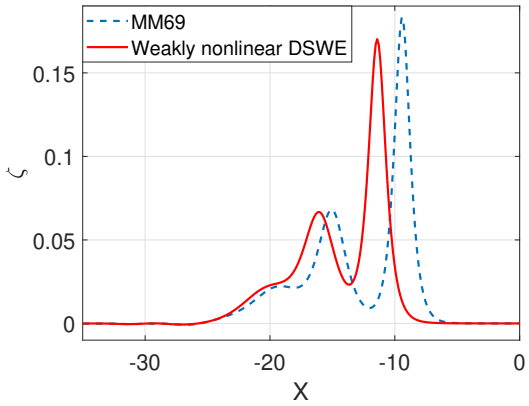


Figure 4.8 Weakly nonlinear system compared with MM69 at $t = 62$.

If the amplitude is increased to $a = 0.3$, we can see a clear difference between the strongly nonlinear DSWE and weakly nonlinear model (2.25)–(2.27), as shown in Figure 4.9. At $t = 42$, the numerical solution of both the DSWE and the weakly nonlinear DSWE show clearly the generation of the second solitary waves, but the two waves are not identical. While the numerical solution of the DSWE needs to be validated with experiments, this indicates that the weakly nonlinear model might not be so reliable when the wave amplitude is not small.

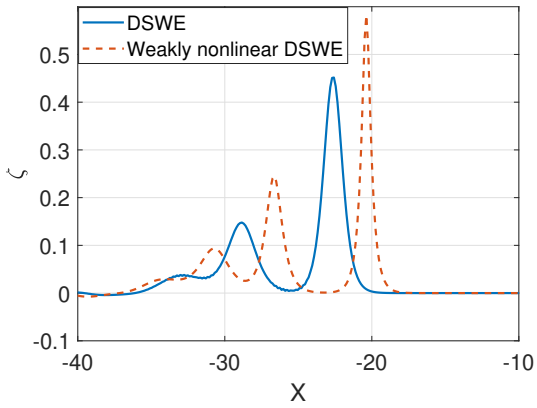


Figure 4.9 Comparision between DSWE and weakly nonlinear system when $t = 42$.

CHAPTER 5

CONCLUSION

In this dissertation, we discussed time-stepping methods for the dispersive shallow water equations with different bathymetry. Depending on the different assumptions of the bathymetry, there are several kinds of simplified DSWEs. After the pseudo-spectral spatial discretization, we recast the DSWE into a differential-algebraic form which can be written as a time evolution ODE plus spatial algebraic constraint. The DAE system can provide a simple form to establish efficient and stable numerical time-stepping schemes.

We provide two time-stepping strategies for solving the DSWE. The first approach is to use a Runge-Kutta method to solve the PDE, and a fixed point iteration scheme with preconditioner to solve the algebraic constraint. The second one is to use an ImEx method, which avoids the need for iterations. When developing the time-stepping method for the dispersive shallow water equations, we need to be careful that stability is ensured for both the Runge-Kutta method with preconditioner, and the ImEx method. The key contribution is discussing how to find an appropriate linear operator that acts as both a preconditioner for the Runge-Kutta method, and also as a splitting for the ImEx method. In both the ImEx and iterative schemes, the choice of $\mathcal{A}$ is constrained by stability. Namely, the (generalized) eigenvalues μ of $\mathcal{A}^{-1}\mathcal{N}$ should be well-conditioned yielding a reasonable conditioning number, and (in the case of the ImEx scheme) also lie inside a stability region.

The ImEx method generally outperforms Runge-Kutta methods of the same order by avoiding the need for iterations. Hence, ImEx methods can shorten computational (clock) time. But, ImEx method treats those terms with higher order derivatives explicitly so that the stability of ImEx schemes should be carefully checked when developing the method.

The dispersive terms in the DSWE lead to physically relevant effects as outlined in Section 4.2. When a solitary wave is propagating over an uneven bottom, the profile of the wave is deformed because of the bathymetry. When there are small amplitude waves, or slowly varying bathymetry (i.e., small grad h), the DSWE reduce to simpler equations that can be used without a loss of efficacy.

APPENDIX

DERIVATION OF DSWE WITH FLAT BOTTOM (D=1)

This appendix focuses on the derivation of the dispersive shallow water equations for one-dimensional waves. Consider the two-dimensional Euler equations with flat bottom ($h(x)$ is any constant number) at this moment. Recast the Euler equations in component form

$$u_t + uu_x + wu_z = -\frac{1}{\rho}P_x, \tag{A.1}$$

$$w_t + uw_x + ww_z = -\frac{1}{\rho}P_z - g, \tag{A.2}$$

with the continuity equation given by

$$u_x + w_z = 0. \tag{A.3}$$

From Figure 2.1, the depth-averaged velocity $\bar{u}(x,t)$ can be introduced as

$$\bar{u}(x,t) = \frac{1}{\zeta + h}\int_{-h}^{\zeta} u(x,z,t)dz. \tag{A.4}$$

Now after integrating equation (A.3) with respect to z from $-h$ to ζ, one can obtain

$$\int_{-h}^{\zeta} u_x dz + \int_{-h}^{\zeta} w_z dz = 0$$

$$\Rightarrow \partial_x \int_{-h}^{\zeta} u dz - u|_{\zeta}\zeta_x + w|_z - w|_{-h} = 0.$$

By applying the boundary condition at the bottom

$$w = 0, \text{ at } z = -h, \tag{A.5}$$

and the kinematic boundary condition at the free surface

$$\zeta_t + u\zeta_x = w, \text{ at } z = \zeta, \tag{A.6}$$

equation (A.3) can be written as

$$\zeta_t + [(h+\zeta)\bar{u}]_x = 0. \tag{A.7}$$

Similarly, integrating equation (A.1) yields

$$\int_{-h}^{\zeta} u_t dz + \int_{-h}^{\zeta} uu_x dz + \int_{-h}^{\zeta} wu_z dz = -\frac{1}{\rho}\int_{-h}^{\zeta} P_x dz$$

$$\begin{aligned} \Rightarrow \partial_t \int_{-h}^{\zeta} u dz - \zeta_t u|_z + \partial_x \int_{-h}^{\zeta} uu dz - \int_{-h}^{\zeta} uu_x dz \\ -\zeta_x uu|_\zeta + uw|_{-h}^{\zeta} - \int_{-h}^{\zeta} uw_z dz = -\frac{1}{\rho}\int_{-h}^{\zeta} P_x dz. \end{aligned}$$

After using the continuity equation and the kinematic boundary conditions given by (A.5)–(A.6), this equation can be written as

$$\partial_t \int_{-h}^{\zeta} u dz + \partial_x \int_{-h}^{\zeta} uu dz = -\frac{1}{\rho}\int_{-h}^{\zeta} P_x dz. \tag{A.8}$$

With the definition of the depth-averaged velocity, equation (A.8) can be rewritten into

$$(\eta\bar{u})_t + (\eta\overline{uu})_x = -\frac{1}{\rho}\eta\overline{P_x}, \tag{A.9}$$

where $\eta = \zeta + h$. To close the system given by (A.7) and (A.9), asymptotic expansion for long waves is introduced to the Euler equations. As introduced in Section 2.1, we assume $\beta \ll 1$ for long waves where $\beta = \bar{h}/\lambda$.

First, the Euler equations need to be non-dimensionalized as

$$x = \lambda x^*,\ z = hz^*,\ t = \frac{\lambda}{c_0}t^*,$$
$$u = c_0 u^*,\ w = \beta c_0 w^*,\ P = \rho c_0^2 P^*,$$

where $c_0 = \sqrt{gh}$. Notice that MM69 systems use the same set of dimensionless variables. By dropping all the $*$ for brevity, the dimensionless Euler equations have the form

$$u_t + uu_x + wu_z = -P_x, \tag{A.10}$$
$$\beta^2(w_t + uw_x + ww_z) = -P_z - 1, \tag{A.11}$$

while the dimensionless continuity equation becomes

$$u_x + w_z = 0. \tag{A.12}$$

For the dimensionless Euler equations, the physical variables are expended as

$$(u, w, P) = (u_0, w_0, P_0) + \beta^2(u_1, w_1, P_1) + O(\beta^4). \tag{A.13}$$

Then the expansions are substituted into equation (A.11), the leading-order equation is found as

$$-P_{0z} - 1 = 0. \tag{A.14}$$

With the boundary condition for the pressure at the free surface $P_0|_{z=\zeta} = 0$, the leading-order pressure P_0 can be found as

$$P_0 = -(z - \zeta). \tag{A.15}$$

Then substitute equation (A.13) into equation (A.10) to find, at $O(1)$:

$$u_{0t} + u_0 u_{0x} + w_0 u_{0z} = -P_{0x}. \tag{A.16}$$

As equation (A.15) leads to $P_{0x} = \zeta_x$, it can be shown that $u_{0z} = 0$ at any time if $u_{0z} = 0$ at $t = 0$. Therefore, we have

$$u_{0t} + u_0 u_{0x} = -\zeta_x, \tag{A.17}$$

$$u_0 = u_0(x, t). \tag{A.18}$$

Then $\overline{uu}$ can be written as

$$\overline{uu} = \overline{u_0 u_0 + O(\beta^2)} = \bar{u}\bar{u} + O(\beta^2),$$

where $\bar{u} = \overline{u_0 + O(\beta^2)} = \overline{u_0} + O(\beta^2)$. This means the depth-averaged equations can be rewritten, after they are non-dimensionalized, as

$$\eta_t + (\eta \bar{u})_x = 0, \tag{A.19}$$

$$\bar{u}_t + \bar{u}\bar{u}_x + \eta_x = O(\beta^2), \tag{A.20}$$

where $\eta = 1 + \zeta$. When $\beta = 0$, equations (A.19)–(A.20) are referred to as the *Shallow Water Equations* with $\eta = \zeta + 1$ (nondimensional form). It can be seen clearly that in the shallow water equations, there are no dispersive terms on the right hand side of (A.20) which will cause shocks after waves traveling for certain time (especially when D=1). To include the dispersive terms to regularize the system, we need to find $O(\beta^2)$ terms from

$$w_{0z} = -u_{0x},$$

$$-P_{1z} = w_{0t} + u_0 w_{0x} + w_0 w_{0z}.$$

After imposing the boundary conditions on the free surface and the bottom ($P_x|_{z=\zeta} = 0$ and $w_0|_{z=-h} = 0$), P_{1x} can be found as

$$P_{1x} = -\frac{1}{3}(\eta^3 G)_x, \tag{A.21}$$

where $G = u_{0xt} + u_0 u_{0xx} - u_{0x}^2$. Then, along with

$$\overline{uu} = \overline{(u_0 + \beta^2 u_1 + O(\beta^4))(u_0 + \beta^2 u_1 + O(\beta^4))} = \bar{u}\bar{u} + O(\beta^4),$$

equations (A.7) and (A.9) can be approximated to

$$\eta_t + (\eta \bar{u})_x = 0, \tag{A.22}$$

$$\bar{u}_t + \bar{u}\bar{u}_x + \eta_x = \frac{\beta^2}{\eta}\left(\frac{\eta^3}{3}G\right)_x + O(\beta^4), \tag{A.23}$$

where $\eta = 1 + \zeta$ (dimensionaless form) and $G = \bar{u}_{xt} + \bar{u}\bar{u}_{xx} - \bar{u}_x^2$. After dropping all bars for brevity, equations (A.22)–(A.23) are referred as the *Dispersive Shallow Water Equations* under the flat bathymetry.

REFERENCES

[1] W. Choi. On Rayleigh Expansion for Nonlinear Long Water Waves. *Journal of Hydrodynamics*, 31:1115–1126, 2019.

[2] W. Choi and R. Camassa. Fully Nonlinear Internal Waves in a Two-Fluid System. *Journal of Fluid Mechanics*, 396:1–36, 1999.

[3] P. Concus and G. H. Golub. Use of Fast Direct Methods for the Efficient Numerical Solution for Nonseparable Elliptic Equations. *SIAM Journal on Numerical Analysis*, 10, 1973.

[4] T. A. Driscoll. A Composite Runge-Kutta Method for the Spectral Solution of Semilinear PDEs. *Journal of Computational Physics*, 182:357–367, 2002.

[5] J. T. Kirby G. Wei and S. T. Grilli. A Fully Nonlinear Boussinesq Model for Surface waves. 1. Highly Nonlinear Unsteady Waves. *Journal of Fluid Mechanics*, 294:71–92, 1995.

[6] A. E. Green and P. M. Naghdi. Derivation of Equations for Wave Propagation in Water of Variable Depth. *Journal of Fluid Mechanics*, 78:237–246, 1976.

[7] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. New York City, NY: Springer, second edition, 1996.

[8] (Lord) J. W. S. Rayleigh. On Waves. *Philosophical Magazine, Series 5*, 1:257–279, 1876.

[9] R. J. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependence Problems*. Philadelphia, PA: SIAM, first edition, 2007.

[10] Y. A. Li, J. M. Hyman, and W. Choi. A Numerical Study of the Exact Evolution Equations for Surface Waves in Water of Finite Depth. *Studies in Applied Mathematics*, 113:303–324, 2004.

[11] O. S. Madsen and C. C. Mei. The Transformation of a Solitary Wave Over an Uneven Bottom. *Journal of Fluid Mechanics*, 39:781–791, 1969.

[12] C. C. Mei. *The Applied Dynamics of Ocean Surface Waves*. River Edge, NJ: World Scientific, first edition, 1989.

[13] J. W. Miles. Solitary Waves. *Annual Review of Fluid Mechanics*, 12:11–43, 1980.

[14] M. Miyata. Long Internal Waves of Large Amplitude. *roceeding of the IUTAM Symposium on Nonlinear Water Waves*, pages 399–406, 1988.

[15] O. Nwogu. An Alternative Form of the Boussinesq Equations for Nearshore Wave Propagation. *Journal of Waterway, Port, Coastal, Ocean Engineering*, 119:618–638, 1993.

[16] H. B. Bingham P. A. Madsen and H. Liu. A New Boussinesq Method for Fully Nonlinear Waves From Shallow to Deep Water. *Journal of Fluid Mechanics*, 462:1–30, 2002.

[17] D. H. Peregrine. Long Waves on a beach. *Journal of Fluid Mechanics*, 27:815–827, 1967.

[18] R. R. Rosalse, B. Seibold, D. Shirokoff, and D. Zhou. Unconditional Stability for Multistep IMEX Schemes: Theory. *SIAM Journal on Numerical Analysis*, 55:2336–2360, 2017.

[19] B. Seibold, D. Shirokoff, and D. Zhou. Unconditional Stability for Multistep IMEX Schemes: Practice. *Journal of Computational Physics*, 376:295–321, 2019.

[20] F. Serre. Contribution à L´étude Des écoulements Permanents et Variables Dans Les Canaux. *La Houille Blanche*, 6:830–872, 1953.

[21] C. H. Su and C. S. Gardner. Korteweg-de Vries Equation and Generalizations. iii: Derivation of the Korteweg-de Vries Equation and Burgers Equation. *Journal of Mathematical Physics*, 10:536–539, 1969.

[22] C. E. Synolakis. The Runup of Solitary Waves. *Journal of Fluid Mechanics*, 185:523–545, 1987.

[23] L. N. Trefethen. *Spectral Methods in Matlab*. Philadelphia, PA: SIAM, 2000.

[24] L. N. Trefethen and III D. Bau. *Numerical Linear Algebra*. Philadelphia, PA: SIAM, 1997.

[25] G. B. Whitham. *Linear and Nonlinear Waves*. Hoboken, NJ: Wiley, 1999.

[26] T. Y. Wu. Modeling Nonlinear Dispersive Water Waves. *Journal of Engineering Mechanics*, 11:747–755, 1999.

[27] T. Y. Wu. A Unified Theory for Modeling Water Waves. *Advances in Applied Mechanics*, 37:1–88, 2001.

[28] P. A. Madsen Y. Agnon and H. A. Schaffer. A New Approach to High-Order Boussinesq Models. *Journal of Fluid Mechanics*, 399:319–333, 1999.