

London Mathematical Society Lecture Note Series: 363

Foundations of Computational Mathematics:
Hong Kong 2008

Edited by

Felipe Cucker
City University of Hong Kong

Allan Pinkus
Technion

Mike Todd
Cornell University

 **CAMBRIDGE**
UNIVERSITY PRESS

Contents

<i>Preface</i>	<i>page</i>	ix
<i>Contributors</i>		xi
1 Smoothed Analysis of Condition Numbers <i>Peter Bürgisser</i>		1
2 A World of Binomials <i>Alicia Dickenstein</i>		42
3 Subdivision Schemes in Geometric Modeling <i>Nira Dyn</i>		68
4 Energy Preserving Schemes <i>Ulrik Fjordholm, Siddhartha Mishra and Eitan Tadmor</i>		93
5 Pathwise Convergence of Numerical Schemes <i>A. Jentzen, P. E. Kloeden and A. Neuenkirch</i>		140
6 Global Behaviour of Conservative Low-Dimensional Systems <i>Carles Simó</i>		162
7 A Panoramic View of Asymptotics <i>R. Wong</i>		190
8 Tractability of Multivariate Problems <i>H. Woźniakowski</i>		236

Preface

The Society for the Foundations of Computational Mathematics supports and promotes fundamental research in computational mathematics and its applications, interpreted in the broadest sense. It fosters interaction among mathematics, computer science and other areas of computational science through its conferences, workshops and publications. As part of this endeavour to promote research across a wide spectrum of subjects concerned with computation, the Society brings together leading researchers working in diverse fields. Major conferences of the Society have been held in Park City (1995), Rio de Janeiro (1997), Oxford (1999), Minneapolis (2002), Santander (2005), and Hong Kong (2008). The next conference is expected to be held in 2011. More information about FoCM is available at its website <http://www.focm.net>.

The conference in Hong Kong on June 16 – 26, 2008, was attended by several hundred scientists. FoCM conferences follow a set pattern: mornings are devoted to plenary talks, while in the afternoon the conference divides into a number of workshops, each devoted to a different theme within the broad theme of foundations of computational mathematics. This structure allows for a very high standard of presentation, while affording endless opportunities for cross-fertilization and communication across subject boundaries. Workshops at the Hong Kong conference were held in the following nineteen fields:

- Approximation theory
- Asymptotic analysis
- Computational algebraic geometry
- Computational dynamics
- Computational number theory
- Foundations of numerical PDEs
- Geometric integration and computational mechanics
- Image and signal processing
- Information-based complexity
- Learning theory
- Multiresolution and adaptivity in numerical PDEs
- Numerical linear algebra

- Optimization
- Real-number complexity
- Relations with computer science
- Special functions and orthogonal polynomials
- Stochastic computation
- Stochastic eigenanalysis
- Symbolic analysis

In addition to the workshops, eighteen plenary lectures, covering a broad spectrum of topics connected to computational mathematics, were delivered by some of the world's foremost researchers. This volume is a collection of articles based on the plenary talks presented at FoCM 2008. The topics covered in the lectures and in this volume reflect the breadth of research within computational mathematics as well as the richness and fertility of interactions between seemingly unrelated branches of pure and applied mathematics.

We hope that this volume will be of interest to researchers in the field of computational mathematics and also to non-experts who wish to gain some insight into the state of the art in this active and significant field.

Like previous FoCM conferences, the Hong Kong gathering proved itself to be a rather uncommon but very stimulating meeting place of researchers in computational mathematics and of theoreticians in mathematics and computer science. While presenting plenary talks by foremost world authorities and maintaining the highest technical level in the workshops, the conference, like previous meetings, laid emphasis on multidisciplinary interaction across subjects and disciplines in an informal and friendly atmosphere.

We wish to express our gratitude to the local organizers and administrative staff of our hosts at the City University of Hong Kong, and wish to thank the Croucher Foundation, the Department of Mathematics at City University of Hong Kong, the Liu Bie Ju Centre for Mathematical Sciences at City University of Hong Kong, and the French General Consulate of Hong Kong and Macau for their financial assistance and for making FoCM 2008 such an outstanding success. We would like to thank the authors of the articles in this volume for producing in short order such excellent contributions. Above all, however, we wish to express our gratitude to all the participants of FoCM 2008 for attending the meeting and making it such an exciting, productive and scientifically stimulating event.

4

Energy Preserving and Energy Stable Schemes for the Shallow Water Equations

Ulrik S. Fjordholm

*Department of Mathematics
University of Oslo*

*P.O. Box 1053, Blindern, N-0316 Oslo, Norway
e-mail: ulriksf@ulrik.uio.no*

Siddhartha Mishra

*Centre of Mathematics for Applications (CMA)
University of Oslo*

*P.O. Box 1053, Blindern, N-0316 Oslo, Norway
e-mail: siddharm@cma.uio.no*

Eitan Tadmor

*Department of Mathematics, Institute for Physical Sciences & Technology
and Center of Scientific Computation and Mathematical Modeling*

*University of Maryland
MD 20742-4015, USA*

e-mail: tadmor@cscamm.umd.edu

Abstract

We design energy preserving and energy stable schemes for the shallow water equations. A new explicit energy preserving flux is proposed and is compared with existing energy preserving fluxes. This new flux results in a considerable reduction in the computational cost. We add suitably discretized viscous terms to the energy preserving scheme in order to obtain an energy stable scheme that replicates the energy decay of the continuous problem. The addition of physical viscosity dramatically reduces the oscillations on resolved meshes. For computing on under-resolved meshes, we propose a new Roe-type numerical flux that adds diffusion in terms of energy variables to an energy preserving scheme. The resulting scheme is energy stable and as accurate as the Roe scheme, at a similar computational cost. The robustness of the energy preserving

and energy stable schemes is demonstrated through several numerical experiments in both one and two space dimensions.

4.1 Introduction

Many interesting models in hydrology and oceanography involve flows where the horizontal scales of motion are much greater than the vertical scale. Such models include flows in lakes, rivers and irrigation channels and near-shore models in oceanography and climate modeling. These phenomena are often modeled by the shallow water equations,

$$\begin{aligned} h_t + (hu)_x + (hv)_y &= 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x + (huv)_y &= \nu((hu_x)_x + (hu_y)_y), \\ (hv)_t + (huv)_x + \left(hv^2 + \frac{1}{2}gh^2\right)_y &= \nu((hv_x)_x + (hv_y)_y), \end{aligned} \quad (4.1)$$

where h is the height of the fluid, u and v are the velocity components in the x and y directions, respectively, g is the constant acceleration due to gravity and ν is the eddy viscosity. The eddy viscosity is responsible for the transfer of energy to the smaller scales of the motion. The system is equipped with suitable initial and boundary conditions.

The eddy viscosity ν generally determines the smallest scale of the flow, and in most applications it is very small. It is common to assume that $\nu = 0$ and consider the inviscid form of the shallow water system,

$$\begin{aligned} h_t + (hu)_x + (hv)_y &= 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x + (huv)_y &= 0, \\ (hv)_t + (huv)_x + \left(hv^2 + \frac{1}{2}gh^2\right)_y &= 0. \end{aligned} \quad (4.2)$$

This is a system of conservation laws, which in general is of the form

$$U_t + F(U)_x + G(U)_y = 0. \quad (4.3)$$

In our case, $U = [h, hu, hv]^\top$ is the vector of conserved quantities of mass and momentum and $F \equiv F(U) = [hu, hu^2 + \frac{1}{2}gh^2, huv]^\top$ and $G \equiv G(U) = [hv, huv, hv^2 + \frac{1}{2}gh^2]^\top$ are the fluxes in the x - and y -directions, respectively. The eigenvalues λ_i and μ_i of the Jacobians F' and G' , respectively, are

$$\lambda_1 = u - \sqrt{gh}, \quad \mu_1 = v - \sqrt{gh},$$

$$\begin{aligned}\lambda_2 &= u, & \mu_2 &= v, \\ \lambda_3 &= u + \sqrt{gh}, & \mu_3 &= v + \sqrt{gh}.\end{aligned}$$

The corresponding eigenvectors are easily calculated; see LeVeque (2002).

Nonlinear systems of conservation laws of the form (4.3) arise in a wide variety of problems in elasticity, fluid dynamics and plasma physics (see Dafermos (2000) for details). The most striking feature about these equations is the fact that even smooth initial data can lead to solutions with discontinuities. This feature is exhibited even in the simplest case of a scalar conservation law in one space dimension. The presence of discontinuities forces us to consider solutions of (4.3) in the weak or averaged sense (see Dafermos (2000) and other standard textbooks for a definition).

Weak solutions are not necessarily unique, and to obtain uniqueness the equations must be augmented with additional admissibility criteria. These criteria assume the form of entropy conditions Dafermos (2000). Many systems of conservation laws are equipped with a convex function $E(U)$ and associated entropy flux functions $H = H(U)$ and $K = K(U)$ such that

$$\partial_{U_i} H(U) = \langle V, \partial_U F^{(i)} \rangle, \quad \partial_{U_i} K(U) = \langle V, \partial_U G^{(i)} \rangle \quad \text{for } i = 1, \dots, n, \quad (4.4)$$

where $V := \partial_U E$ is the vector of *entropy variables*. As an immediate consequence of these identities, smooth solutions of (4.3) will satisfy the additional conservation law

$$E(U)_t + H(U)_x + K(U)_y = 0. \quad (4.5)$$

However, this identity is not valid at shocks and has to be modified accordingly. The entropy identity (4.5) transforms into the entropy *inequality* Dafermos (2000)

$$E(U)_t + H(U)_x + K(U)_y \leq 0, \quad (4.6)$$

in the sense of distributions. Scalar conservation laws are equipped with an infinite number of entropy/entropy-flux pairs, and this paves the way for a proof of existence, uniqueness and stability. However, systems of conservation laws in general do not possess an infinite number of entropy functions. This is a key difficulty in proving existence and stability results, particularly in the case of multi-dimensional systems.

However, many interesting systems like the Euler equations of gas dynamics and the magnetohydrodynamics (MHD) system of plasma

physics are equipped with at least one physically relevant entropy function. The shallow water system (4.2) also possesses an entropy function, the *total energy*

$$E = \frac{1}{2} (hu^2 + hv^2 + gh^2),$$

which is the sum of kinetic and gravitational potential energy. A direct calculation reveals that smooth solutions of (4.2) satisfy the energy conservation law

$$E_t + \left(\frac{1}{2} (hu^3 + huv^2) + gh^2 \right)_x + \left(\frac{1}{2} (hu^2v + hv^3) + gh^2 \right)_y = 0. \quad (4.7)$$

Integrating (4.7) over space, we obtain

$$\frac{d}{dt} \int_{\mathbb{R}^2} E \equiv 0. \quad (4.8)$$

Hence, the total energy of smooth solutions of (4.2) is conserved.

The above identity is valid only for smooth solutions of (4.2). Energy will be dissipated at shocks, and the precise rate of this dissipation can be explicitly calculated from the viscous form (4.1). The energy identity (4.7) in the presence of viscous terms takes the form,

$$\begin{aligned} E_t + \left(\frac{1}{2} (hu^3 + huv^2) + gh^2 \right)_x + \left(\frac{1}{2} (hu^2v + hv^3) + gh^2 \right)_y \\ = \nu \left(u((hu_x)_x + (hu_y)_y) + v((hv_x)_x + (hv_y)_y) \right). \end{aligned} \quad (4.9)$$

Integrating this identity in space and integrating by parts we obtain the energy dissipation estimate

$$\frac{d}{dt} \int_{\mathbb{R}^2} E = -\nu \int_{\mathbb{R}^2} h (u_x^2 + u_y^2 + v_x^2 + v_y^2). \quad (4.10)$$

As the height h is always positive, the right-hand side of the above identity is always non-positive, so we get the energy dissipation

$$\frac{d}{dt} \int_{\mathbb{R}^2} E \leq 0. \quad (4.11)$$

Thus, we recover the energy inequality that holds for weak solutions of (4.2). Furthermore, (4.10) gives an explicit rate for the dissipation of energy into smaller scales. Note that a bound on the total energy E automatically implies a bound on the L^2 norms of height and the velocity field, as the height is strictly positive. Hence, the energy estimate is also a statement of stability of solutions of the shallow water system.

However, energy estimates are not enough to obtain proofs of existence or uniqueness of solutions.

In the absence of existence results or explicit formulas for the solution of (4.3), numerical methods are the main tools in the study of these models. Numerical methods for systems of conservation laws have undergone extensive development in the last few decades, and the subject is in a fairly mature stage; see LeVeque (2002). The most popular methods are the so-called finite volume methods. For simplicity, we consider a uniform Cartesian mesh in \mathbb{R}^2 with mesh sizes Δx and Δy , respectively. We denote the nodes as $x_i := i\Delta x$ and $y_j := j\Delta y$ and a prototypical cell as $I_{ij} := [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}) \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}})$. A standard cell-centered finite volume method consists of updating the cell averages,

$$U_{ij}(t) := \frac{1}{\Delta x \Delta y} \int_{I_{ij}} U(x, y, t) dx dy,$$

at each time level. For simplicity, we drop the time dependence of every quantity and write a standard finite volume scheme for (4.3) in the semi-discrete form as

$$\frac{d}{dt} U_{ij} = -\frac{1}{\Delta x} \left(F_{i+\frac{1}{2}, j} - F_{i-\frac{1}{2}, j} \right) - \frac{1}{\Delta y} \left(G_{i, j+\frac{1}{2}} - G_{i, j-\frac{1}{2}} \right), \quad (4.12)$$

where $F_{i\pm\frac{1}{2}, j\pm\frac{1}{2}}$ and $G_{i\pm\frac{1}{2}, j\pm\frac{1}{2}}$ are numerical fluxes at the cell-edges, consistent with the fluxes F and G , respectively¹. The key step is the choice of the numerical fluxes. The numerical fluxes are computed using the neighboring cell-averages across the normal directions of cell-edges, and high-order accuracy in space is obtained with non-oscillatory reconstructions of point values from these cell-averages, e.g., Cockburn et al. (1998), LeVeque (2002) and the references therein. The time-integration is often performed with strong stability preserving Runge-Kutta methods, e.g., Gottlieb, Shu and Tadmor (2001). This standard framework has proved very successful in computing solutions of many interesting flow problems and is used extensively in practice.

However, very few rigorous stability results are obtained for finite volume schemes, particularly for systems of conservation laws. One of the reasons for the lack of stability is the failure to design schemes that satisfy a discrete form of the entropy inequality (4.6). Most finite volume

¹ Note that the differential fluxes $F(U), G(U)$ depend on the one conserved quantity U whereas numerical fluxes depend on two or more neighboring cell quantities, e.g., $F_{i+\frac{1}{2}, j} = F(\dots, U_{ij}, U_{i+1, j}, \dots)$. This enables us to uniquely distinguish between differential fluxes such as $F_{ij} = F(U_{ij})$, and the corresponding numerical fluxes, e.g., $F_{i\pm\frac{1}{2}, j}$ etc.

schemes (particularly at higher-order) do not necessarily satisfy such inequalities. This might result in numerical instabilities and the computation of incorrect solutions. Many finite volume schemes (in their first-order versions) do add enough numerical diffusion to dissipate entropy; however, the diffusion added is excessive and does not respect the rate of entropy diffusion in the continuous problem.

In the specific context of the shallow water equations (4.2), there are many different finite volume schemes in use. Most of these schemes do not preserve energy for smooth solutions (i.e., satisfy discrete versions of (4.7)), which can lead to significant loss of accuracy, particularly for problems involving large time scales. Few schemes will respect the energy balance (4.10) because entropy dissipation at shocks is too excessive. This can lead to both instabilities as well as numerical artifacts. See Arakawa (1966), Arakawa and Lamb (1977), Arakawa and Lamb (1981) for a detailed discussion on numerical effects of schemes that do not respect the energy balance in shallow water equations.

In the pioneering papers of Tadmor (1987) and (2003), the problem of *designing* finite volumes schemes which satisfy discrete versions of the entropy inequality (4.6) was tackled. The main feature of Tadmor (1987) was the design of an *entropy conservative* finite volume flux, i.e., consistent numerical fluxes $\tilde{F}_{i\pm\frac{1}{2},j}$ and $\tilde{G}_{i,j\pm\frac{1}{2}}$, which ensure that the numerical scheme (4.12) satisfies the entropy identity (4.5). Then, a novel entropy comparison principle is introduced in order to design an *entropy stable* scheme – a scheme that satisfies a discrete version of the entropy inequality. The idea is to compare the numerical diffusion of any given finite volume scheme with the diffusion of the entropy conservative scheme. A scheme that contains more diffusion than an entropy conservative scheme is entropy stable. Thus, one is able to investigate entropy stability using a comparison principle. Tadmor (1987) did not, however, contain explicit expressions for any interesting systems of conservation laws. A novel pathwise decomposition was introduced in Tadmor (2003) in order to obtain an explicit formula for the entropy conservative scheme. This approach was used in Tadmor and Zhong (2006) to compute solutions of the Euler equations, and in Tadmor and Zhong (2008) to approximate the shallow water equations. In both papers the authors used the explicit entropy conservative scheme of Tadmor (2003) to compute the solutions and used this scheme as a basis for a “faithful” discretization of the entropy (energy) balance.

Our aim in this paper is to consider energy preserving and energy stable discretizations of the shallow water system. We consider three

different energy preserving finite volume schemes: The original entropy conservative scheme of Tadmor (1987), which can be explicitly integrated in the case of shallow water equations in one dimension; the pathwise explicit scheme of Tadmor (2003), Tadmor and Zhong (2008); and a novel explicit energy preserving scheme. We compare the three energy preserving schemes, and find that the new explicit entropy preserving scheme is both simpler to implement and computationally less expensive compared to the other two schemes. We use this new energy conservative scheme to design novel numerical diffusion operators that result in energy stable schemes. The schemes are implemented in a series of numerical experiments which illustrate their different features in the one-dimensional setup, in Section 4.2 and with the two-dimensional problems, in Section 4.3. Conclusions are drawn in Section 4.4.

4.2 The One-Dimensional Problem

For simplicity of the description, we start with the one-dimensional form of the inviscid shallow water system (4.2),

$$\begin{aligned} h_t + (hu)_x &= 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x &= 0. \end{aligned} \quad (4.13)$$

These equations are obtained from (4.2) simply by ignoring variation in the y -direction and setting the vertical velocity component $v = 0$.

The above equation is a form of the generic one-dimensional system of conservation laws,

$$U_t + F_x = 0, \quad (4.14)$$

with U the n -vector of unknowns and $F = F(U)$ the flux vector. Assume that an entropy/entropy-flux function pair (E, H) exists, so that

$$E(U)_t + H(U)_x = 0. \quad (4.15)$$

Define the vector of entropy variables as $V := \partial_U E$. For the one-dimensional shallow water system, the entropy function is given by the *energy*, $E = \frac{1}{2}(hu^2 + gh^2)$, which for smooth solutions satisfies

$$E_t + \left(\frac{1}{2}hu^3 + gh^2\right)_x = 0. \quad (4.16)$$

The vector V takes the form $V = \left[gh - \frac{u^2}{2}, u\right]^T$. Define the *entropy*

potential as $\Psi := \langle V, F \rangle - H$. A direct calculation shows that for the one-dimensional shallow water equations, Ψ is given by $\Psi = \frac{1}{2}guh^2$.

Our aim is to design a finite volume scheme for (4.14) that satisfies a discrete form of the entropy identity (4.15). A finite volume scheme (in semi-discrete form) on a uniform mesh $x_i = i\Delta x$ is given by,

$$\frac{d}{dt}U_i = -\frac{1}{\Delta x}(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}), \quad (4.17)$$

where U_i is the cell average on $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ and $F_{i+\frac{1}{2}}$ is the numerical flux at the interface $x_{i+\frac{1}{2}}$.

As mentioned in the introduction, an arbitrary choice of a consistent numerical flux is not enough to satisfy a discrete version of the entropy identity (4.15). Instead, we follow the general procedure introduced in Tadmor (1987) to define an entropy preserving numerical flux. Here and below we use the following abbreviations

$$\begin{aligned} V_i &= V(U_i), & F_i &= F(U_i), & \Psi_i &= \Psi(U_i), \\ \llbracket a_{i+\frac{1}{2}} \rrbracket &:= a_{i+1} - a_i, & \bar{a}_{i+\frac{1}{2}} &:= \frac{1}{2}(a_i + a_{i+1}), \end{aligned}$$

where $\llbracket a_{i+\frac{1}{2}} \rrbracket$ represents the jump of a across the interface at $x_{i+\frac{1}{2}}$.

Theorem 4.1 (Tadmor (1987)) *Consider the one-dimensional system of conservation laws (4.14) with entropy function E , entropy variables V , entropy flux F and potential Ψ , as defined above. Let $\tilde{F}_{i+\frac{1}{2}}$ be a numerical flux, consistent with (4.14), that satisfies*

$$\langle \llbracket V_{i+\frac{1}{2}} \rrbracket, \tilde{F}_{i+\frac{1}{2}} \rangle = \llbracket \Psi_{i+\frac{1}{2}} \rrbracket. \quad (4.18)$$

Then, the scheme (4.17) satisfies the discrete entropy identity

$$\frac{d}{dt}E(U_i(t)) = -\frac{1}{\Delta x}(\tilde{H}_{i+\frac{1}{2}} - \tilde{H}_{i-\frac{1}{2}}), \quad (4.19)$$

with numerical entropy flux $\tilde{H}_{i+\frac{1}{2}} := \langle \bar{V}_{i+\frac{1}{2}}, \tilde{F}_{i+\frac{1}{2}} \rangle - \bar{\Psi}_{i+\frac{1}{2}}$. Summing over all i , we end up with the conservation law $\frac{d}{dt} \sum_i E(U_i(t)) \equiv 0$.

Hence, the finite volume scheme (4.17), (4.18) is energy preserving.

The proof of the above theorem can be found in Tadmor (1987). The theorem is very general and the key ingredient for obtaining entropy preserving fluxes is the condition (4.18).

Note that the condition (4.18) provides a single constraint for a flux to be entropy conservative. We will reserve the notation of \tilde{F} to distinguish

such entropy conservative fluxes. Thus, for example, when $n = 1$ entropy conservative fluxes are uniquely determined as $\tilde{F}_{i+\frac{1}{2}} = [\Psi_{i+\frac{1}{2}}]/[V_{i+\frac{1}{2}}]$. For general $n \times n$ systems, however, the choice of an entropy preserving flux in (4.14) is not unique for $n > 1$. Given this fact, we proceed to describe three different choices of entropy conservative fluxes for the 1D shallow water system (4.13).

4.2.1 Scheme I: Averaged Energy Conservative (AEC) Scheme

We start with an entropy conservative scheme that was first proposed in Tadmor (1987) for a general system (4.14). For $\xi \in [-1/2, 1/2]$, define the straight line

$$V_{i+\frac{1}{2}}(\xi) = \frac{1}{2}(V_i + V_{i+1}) + \xi(V_{i+1} - V_i).$$

Clearly, $V_{i+\frac{1}{2}}$ connects the vectors V_i and V_{i+1} . Next, we define the numerical flux as

$$\tilde{F}_{i+\frac{1}{2}} = \int_{-1/2}^{1/2} F(V_{i+\frac{1}{2}}(\xi)) d\xi. \quad (4.20)$$

This flux represents the path integral of the flux along a straight line connecting two adjacent states in phase space. Clearly, this flux is consistent and it was shown in Tadmor (1987) that it is also entropy conservative. This follows from a straightforward calculation showing that (4.20) satisfies the identity (4.18).

In general, (4.20) cannot be evaluated explicitly in the phase space. But we will see that for the simple case of shallow water equations, the path integral can be explicitly computed in terms of the physical variables although the resulting formulas are still quite complicated. A long direct calculation of the integral in (4.20) for the shallow water system yields the following formulas:

$$\begin{aligned} F_{i+\frac{1}{2}}^{(1)} &= \frac{h_i u_i}{3} + \frac{h_{i+1} u_{i+1}}{3} + \frac{h_i u_{i+1}}{6} + \frac{h_{i+1} u_i}{6} \\ &\quad - \frac{u_i^3}{24} - \frac{u_{i+1}^3}{24} + \frac{u_i u_{i+1}^2}{24} + \frac{u_{i+1} u_i^2}{24} \\ F_{i+\frac{1}{2}}^{(2)} &= \frac{1}{12} h_i u_i^2 + \frac{1}{12} h_{i+1} u_{i+1}^2 + \frac{1}{6} h_i u_{i+1}^2 + \frac{1}{6} h_{i+1} u_i^2 \\ &\quad + \frac{1}{4} h_i u_i u_{i+1} + \frac{1}{4} h_{i+1} u_i u_{i+1} + \frac{7g}{24} h_i^2 + \frac{7g}{24} h_{i+1}^2 \\ &\quad - \frac{g}{12} h_i h_{i+1} + \frac{1}{96} u_i^4 + \frac{1}{96} u_{i+1}^4 - \frac{1}{48} u_i^2 u_{i+1}^2. \end{aligned} \quad (4.21)$$

A direct calculation verifies that this flux is indeed energy preserving. Note that the flux is symmetric in its arguments, and it is easy to check that it is consistent. The main problem with this flux is the complexity of the formulas: This 1D computation is rather specific and a similar calculation to obtain an explicit form of (4.20) for the two-dimensional shallow water system (4.2) becomes algebraically intractable. This necessitates the search for alternative fluxes satisfying (4.18).

4.2.2 Scheme II: Pathwise Energy Conservative (PEC) Scheme

An explicit solution of (4.18) was found in Tadmor (2003). This scheme was implemented for the shallow water system (4.2) in Tadmor and Zhong (2008), and we mention it here simply for the sake of completeness and comparison.

Consider the $n \times n$ system of conservation laws (4.14). Let $\{r_k\}, \{l_k\}$ for $k = 1, \dots, n$ be any orthogonal eigensystem that spans \mathbb{R}^n . At an interface $x_{i+\frac{1}{2}}$, we have the two adjacent entropy variable vectors $V^{[0]} := V_i$ and $V^{[n]} := V_{i+1}$. Then, define the following paths:

$$\begin{aligned} V^{[0]} &= V_i, \\ V^{[k]} &= V^{[k-1]} + \langle [V_{j+\frac{1}{2}}], l_k \rangle r_k \quad \text{for } k = 1, \dots, n. \end{aligned}$$

Note that $V^{[n]} = V_{i+1}$. We are replacing the straight line joining the two adjacent states in the flux (4.20) by a piecewise linear path that corresponds to the basis vectors. Now define, for $k = 1, 2, \dots, n$,

$$\tilde{F}^{[k]} = \frac{\Psi(V^{[k]}) - \Psi(V^{[k-1]})}{\langle [V_{j+\frac{1}{2}}], l_k \rangle} l_k. \quad (4.22)$$

Then the PEC flux is given by

$$\tilde{F}_{i+\frac{1}{2}} = \sum_{k=1}^n \tilde{F}^{[k]}. \quad (4.23)$$

To show that this flux is entropy conservative, multiply both sides of (4.23) with $[V_{i+\frac{1}{2}}]$ to get

$$\begin{aligned} \langle [V_{i+\frac{1}{2}}], \tilde{F}_{i+\frac{1}{2}} \rangle &= \sum_{k=1}^n \left(\Psi(V^{[k]}) - \Psi(V^{[k-1]}) \right) \\ &= \Psi(V^{[n]}) - \Psi(V^{[0]}) = [V_{i+\frac{1}{2}}]. \end{aligned}$$

Consistency of this flux has been shown in Tadmor (2003). The only remaining step in designing the flux is to specify the choice of the path, i.e.,

of the orthogonal eigensystem. Following Tadmor and Zhong (2008) and (2006), we take the path given by the eigenvectors of a Roe matrix corresponding to the shallow water system. Details of the implementation are provided in Tadmor and Zhong (2006). A significant modification is made in Remark 3.5 of Tadmor and Zhong (2006) to treat the case where $\langle [V_{j+\frac{1}{2}}], l_k \rangle$ vanish, making the flux $\tilde{F}^{[k]}$ in (4.22) singular; this case has been treated in Tadmor and Zhong (2006) and (2008).

The main feature of the flux (4.23) is its generality. It provides a recipe to construct entropy conservative schemes for any system. The main difficulty, though, is its computational cost, which is high because one has to evaluate the eigensystem and solve for the orthogonal system for every mesh point. Even though it has been implemented for both the shallow water and Euler equations, we would like to find a energy preserving flux which is computationally cheaper and easier to implement.

4.2.3 Scheme III: Explicit Energy Conservative (EEC) Scheme

Both of the above schemes can be explicitly written down and implemented, although at a high computational cost. Our aim is to design a much simpler scheme. To construct the entropy conservative flux at the interface $x_{i+\frac{1}{2}}$ we will use the following identity between jumps and averages,

$$[a_{i+\frac{1}{2}} b_{i+\frac{1}{2}}] \equiv \bar{b}_{i+\frac{1}{2}} [a_{i+\frac{1}{2}}] + [b_{i+\frac{1}{2}}] \bar{a}_{i+\frac{1}{2}}. \quad (4.24)$$

In order to satisfy the entropy preserving constraint (4.18) for the special case of the shallow water equations (4.13) we use (4.24) to express the jumps across $x_{i+\frac{1}{2}}$ in terms of the jumps in the primitive variables h and u . We have

$$\begin{aligned} [V_{i+\frac{1}{2}}^{(1)}] &= [gh_{i+\frac{1}{2}} - \frac{1}{2}u_{i+\frac{1}{2}}^2] = g[h_{i+\frac{1}{2}}] - \bar{u}_{i+\frac{1}{2}}[u_{i+\frac{1}{2}}], \\ [V_{i+\frac{1}{2}}^{(2)}] &= [u_{i+\frac{1}{2}}], \\ [\Psi_{i+\frac{1}{2}}] &= \frac{1}{2}g[u_{i+\frac{1}{2}}h_{i+\frac{1}{2}}^2] = g\bar{u}_{i+\frac{1}{2}}\bar{h}_{i+\frac{1}{2}}[h_{i+\frac{1}{2}}] + \frac{g}{2}(\bar{h}^2)_{i+\frac{1}{2}}[u_{i+\frac{1}{2}}]. \end{aligned}$$

Writing down the desired flux componentwise as $\tilde{F}_{i+\frac{1}{2}} = [\tilde{F}_{i+\frac{1}{2}}^{(1)}, \tilde{F}_{i+\frac{1}{2}}^{(2)}]$, inserting all the above quantities into (4.18) and then equating jumps

in h and u , we get the following set of equations:

$$\begin{aligned}\tilde{F}_{i+\frac{1}{2}}^{(1)} &= \bar{h}_{i+\frac{1}{2}} \bar{u}_{i+\frac{1}{2}}, \\ \tilde{F}_{i+\frac{1}{2}}^{(2)} - \bar{u}_{i+\frac{1}{2}} \tilde{F}_{i+\frac{1}{2}}^{(1)} &= \frac{g}{2} (\bar{h}^2)_{i+\frac{1}{2}}.\end{aligned}$$

Solving the above equations, we get

$$\begin{aligned}\tilde{F}_{i+\frac{1}{2}}^{(1)} &= \bar{h}_{i+\frac{1}{2}} \bar{u}_{i+\frac{1}{2}}, \\ \tilde{F}_{i+\frac{1}{2}}^{(2)} &= \bar{h}_{i+\frac{1}{2}} \left(\bar{u}_{i+\frac{1}{2}} \right)^2 + \frac{g}{2} (\bar{h}^2)_{i+\frac{1}{2}}.\end{aligned}\tag{4.25}$$

Thus, we can write down the energy preserving flux explicitly by expanding (4.25) as

$$\begin{aligned}\tilde{F}_{i+\frac{1}{2}}^{(1)} &= \left(\frac{h_i + h_{i+1}}{2} \right) \left(\frac{u_i + u_{i+1}}{2} \right), \\ \tilde{F}_{i+\frac{1}{2}}^{(2)} &= \left(\frac{h_i + h_{i+1}}{2} \right) \left(\frac{u_i + u_{i+1}}{2} \right)^2 + \frac{g}{2} (h_i^2 + h_{i+1}^2).\end{aligned}\tag{4.26}$$

Clearly the flux (4.26) is energy preserving as well as consistent. It is symmetric and second-order accurate in space (as shown by a simple truncation error analysis). It is also extremely easy to code. Furthermore, numerical experiments will reveal that the flux is very cheap computationally and is robust. The above flux is similar in spirit, though different in details, to the entropy conservative flux for the Euler equations of gas dynamics designed in Roe (2006).

Remark 4.1 Both the AEC scheme (4.21) and the PEC scheme (4.23) were based on integrating the flux along a suitable path in the space of energy variables. The AEC scheme relied on a straight line path connecting the adjacent states whereas the PEC scheme was based on a piecewise straight line path parallel to eigenvectors of the Jacobian of the flux. A natural question arises — can the EEC scheme (4.26) be written as an energy preserving flux using integration along a suitable path in the phase space of energy variables? We were unable to obtain such a path; however, we believe it exists, and if so, it would be extremely interesting if it can be written down explicitly.

4.2.4 Time Stepping

The three energy conservative schemes above were formulated in the semi-discrete framework (4.17). We also need to describe the discrete

time evolution in order to compute approximations to (4.13) and to this end, we choose two different time stepping schemes. First, note that all the three energy preserving schemes are central schemes and hence unstable when used together with the forward Euler method. Therefore, we need to use Runge-Kutta methods to stabilize the computations. Furthermore, it is advisable to use the strong-stability preserving (SSP) Runge-Kutta methods developed in Gottlieb, Shu and Tadmor (2001). Given a numerical flux F , let $\mathcal{L}(U_i^n)$ denote the net flux into grid cell i at time t_n , i.e.,

$$\mathcal{L}(U_i^n) := -\frac{1}{\Delta x} \left(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} \right).$$

$\mathcal{L}(U_i^n)$ is precisely the right-hand side of the finite volume scheme (4.17). We use the following second-order SSP Runge-Kutta method of Gottlieb, Shu and Tadmor (2001), identified below as RK2:

$$\begin{aligned} U_i^* &= U_i^n + \Delta t \mathcal{L}(U_i^n) \\ U_i^{**} &= U_i^* + \Delta t \mathcal{L}(U_i^*) \\ U_i^{n+1} &= \frac{1}{2}(U_i^n + U_i^{**}), \end{aligned} \tag{RK2}$$

and the third-order SSP Runge-Kutta method of Gottlieb, Shu and Tadmor (2001), denoted RK3:

$$\begin{aligned} U_i^* &= U_i^n + \Delta t \mathcal{L}(U_i^n) \\ U_i^{**} &= \frac{3}{4}U_i^n + \frac{1}{4}U_i^* + \frac{\Delta t}{4}\mathcal{L}(U_i^*) \\ U_i^{n+1} &= \frac{1}{3}U_i^n + \frac{2}{3}U_i^{**} + \frac{2\Delta t}{3}\mathcal{L}(U_i^{**}). \end{aligned} \tag{RK3}$$

4.2.5 Energy Preserving Schemes: Numerical Experiments 1–2

We have three different energy preserving schemes for the shallow water system. We denote the finite volume scheme (4.17) with the energy conservative fluxes (4.21), (4.23) and (4.26) as AEC, PEC and EEC schemes, respectively. Our aim in this section is to compare these three schemes on a series of numerical experiments in one space dimension.

Numerical experiment #1: one-dimensional dam break

We start with a standard one-dimensional dam-break Riemann problem with initial data,

$$h(x, 0) = \begin{cases} 2 & \text{if } x < 0, \\ 1.5 & \text{if } x > 0, \end{cases} \quad u(x, 0) \equiv 0. \quad (4.27)$$

In this problem, the initial data itself is discontinuous, and the exact solution consists of a left going rarefaction and a right going shock. We set the acceleration due to gravity to be $g = 1$. The computational domain is $[-1, 1]$ with 100 mesh points. We use RK2 time stepping scheme at a CFL number of 0.45, and we compute up to time $t = 0.4$. All the computations are performed with transparent Neumann type boundary conditions.

To begin with, we show the heights computed with the AEC, PEC and EEC schemes. The results are shown in Figure 4.1. In this figure, we also show a time history of the energy computed with all the three schemes.

As shown in Figure 4.1, all the three schemes are capturing the essential properties of the correct solution. The dam-break problem has a shock, and energy should be dissipated at a shock. Since the AEC, EEC and PEC schemes are energy preserving, the energy dissipation at the shock does not take place. Instead, the energy preserving schemes simply redistribute the energy into smaller scales in the form of oscillations trailing the shocks. This behavior is standard for energy (entropy) conservative schemes; see Tadmor and Zhong (2008) and the references therein. There is very little difference between the computed solutions.

Next we compute the behavior of the energy as a function of time. Figure 4.1 shows growth of energy of order 10^{-4} for all three schemes. The energy plots of the EEC and PEC schemes lie on top of each other, while the AEC scheme produces slightly more energy – at least initially. This small-magnitude energy *growth*, which may appear puzzling at first sight, is solely due to time discretization. In fact, an RK2 time discretization of entropy conservative scheme will *produce* energy of order $\mathcal{O}(\Delta t)^3$, which explains the growth of energy as seen in Figure 4.1. We note in passing that fully discrete entropy conservative time-integration schemes are discussed in Tadmor (2003), LeFloch, Mercier and Rohde (2002).

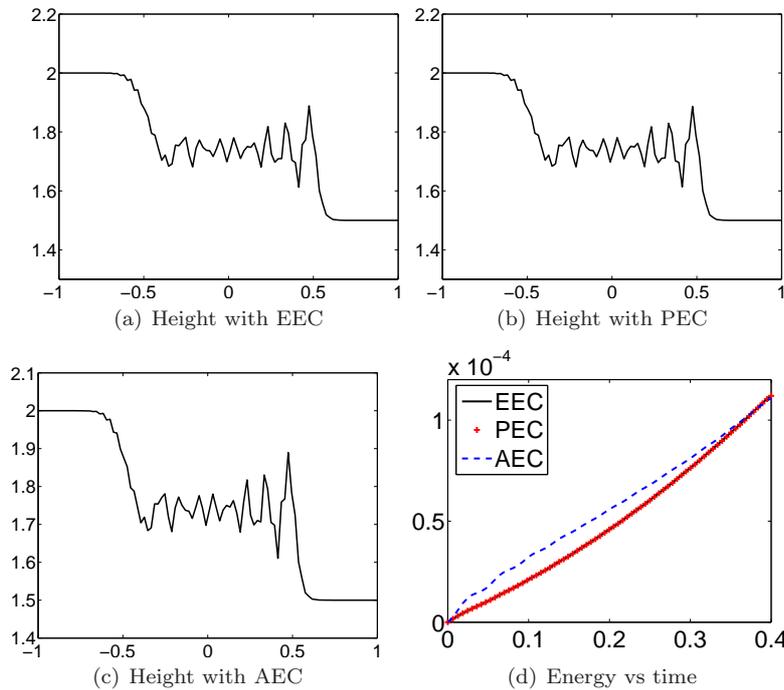


Fig. 4.1. Height h at $t = 0.4$ computed with the three energy conservative schemes on 100 mesh points with an RK2 method and a time history of energy computed with all the three schemes.

Effect of time stepping

The above comments need to be verified with more computations and variation of the time stepping routine and the CFL number. We therefore compare the semi-discrete EEC scheme which is discretized by the second- and third-order Runge-Kutta time stepping, RK2 and RK3, computed with CFL numbers 0.45 and 0.05. The energy history is computed and shown in Figure 4.2. This figure shows the effect time stepping schemes have on the energy balance. Comparing the RK2 results with two different CFL numbers reflects energy growth of order $\mathcal{O}(\Delta t)^3$. Thus, the energy growth of $\sim 10^{-7}$ corresponding to a CFL number 0.05 is three orders of magnitude smaller than the energy growth of $\sim 10^{-4}$ for a CFL number 0.45.

Next, we discuss the effect of using a higher-order time integration scheme. Observe that the third-order RK3 actually *dissipates* energy.

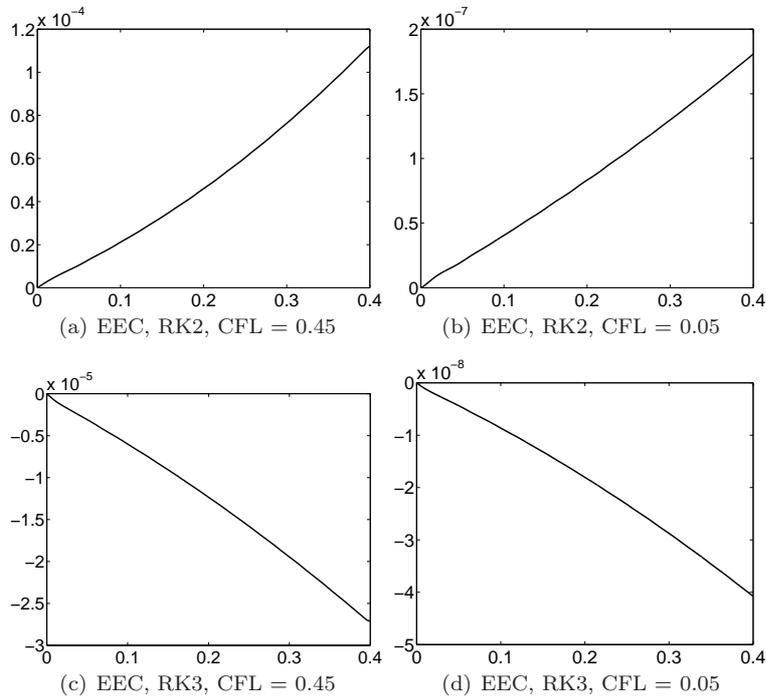


Fig. 4.2. A time history of energy E computed with the EEC scheme on 100 mesh points for two different RK schemes at two different CFL numbers.

Indeed, consideration of absolute stability regions implies that one needs at least third-order discretizations to enforce energy dissipation of entropy conservative scheme, e.g., Tadmor (2002). Moreover, energy dissipates at a rate of order $\mathcal{O}(\Delta t)^3$, ranging from $\sim 10^{-5}$ with CFL = 0.45 to $\sim 10^{-8}$ with CFL = 0.05. Similar results, which we omit, were found for the PEC and AEC schemes.

Effect of resolution: Numerical experiment #2

Another issue of interest for energy preserving schemes is the nature of their dispersive oscillations. To this end we compute the EEC-RK2 scheme with CFL number of 0.45 on four different meshes using 100, 400, 800 and 1600 mesh points, respectively. As shown in Figure 4.3, the solution becomes more and more oscillatory as the mesh is refined. Regardless of mesh size, the amplitude of the oscillations remains bounded and is of the order of the jump in the initial height. However, the fre-

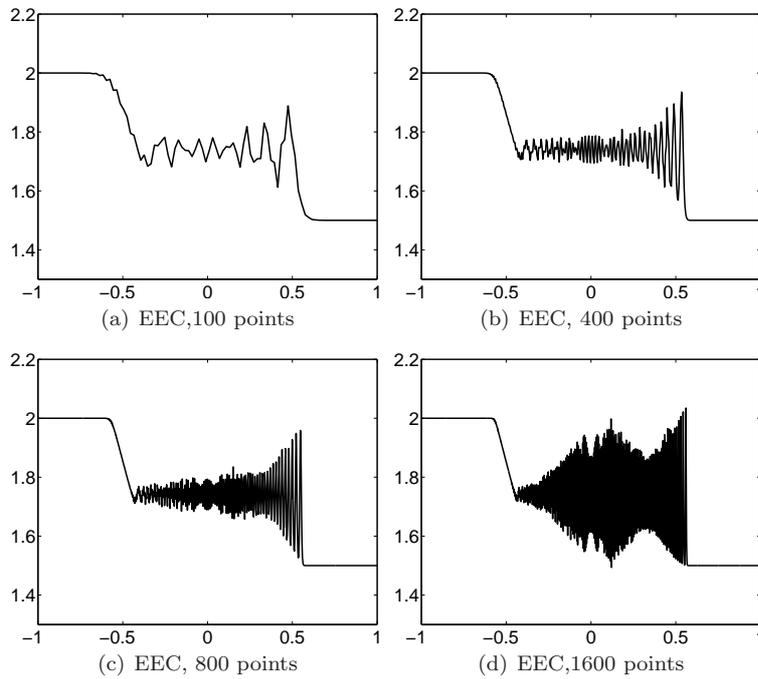


Fig. 4.3. Height h at time $t = 0.4$ computed with EEC scheme and RK2 at CFL 0.45 on four different meshes.

quency of the oscillations keeps increasing until a well-defined wave train is formed behind the shock. In fact, the frequency of the oscillations is at the scale of the mesh size. This wave-train is often referred to as a *modulation*, and is an interesting object of study, see e.g., Goodman and Lax (1988), Hou and Lax (1991) for more details). This behavior is expected as energy, which must be dissipated at shocks, is trapped by our energy preserving scheme, and non-linear dispersive effects redistributes this energy into the smallest resolvable scale on the computational grid, in the form of modulated high-frequencies.

Computational cost

Given the similarities in the behavior of the three energy preserving schemes, a key point is their computational cost. To compare the cost of these schemes we consider the above problem on a fixed mesh of 100 grid points and use the RK2 method of time integration. Then, we change the time step in three different computations so that the energy

change is about 10^{-3} , 10^{-4} and 10^{-5} , respectively. We have attempted to optimize all the schemes as much as possible to obtain a fair comparison of run-times. All the three schemes were implemented in C++ using the Blitz++ numerical linear algebra package. The PEC scheme in particular needs to be carefully implemented to obtain optimized run-times.

Energy error	10^{-3}	10^{-4}	10^{-5}
EEC	1	1.69	3.49
PEC	2.61	4.68	10.47
AEC	1.14	1.91	-

Table 4.1. *Normalized run-times for the three energy preserving schemes on the one-dimensional dam-break problem with three different levels of error in energy.*

In Table 4.1, we show the normalized run-times with the three different schemes. The times are normalized in units where 0.0126 seconds represents unity. As shown in this table, the EEC scheme is the fastest. In fact, the PEC scheme is about three times slower than the EEC scheme for the same level of energy error. The AEC scheme is about 1.2 times slower than the EEC scheme. We were unable to obtain a time step small enough so that the AEC scheme gave an energy error of about 10^{-5} .

Summarizing, we see that all the three energy preserving schemes preserve energy to a satisfactory level. The energy dissipation/production due to explicit time stepping can be reduced by using smaller time steps and higher-order time integration methods. The presence of shocks in the computed solution results in high-frequency oscillations as energy is distributed into smaller scales. In terms of computational cost, the EEC scheme is the best, being about three times cheaper than the PEC scheme. This is expected as the PEC requires eigenvector decompositions at each cell interface. The EEC scheme, on the other hand, is explicit and requires only a few simple floating point operations.

4.2.6 Eddy Viscosity: Numerical Experiment 3

We obtain the one-dimensional form of the shallow water equations with eddy viscosity (4.1) by setting v and all change in y -direction to zero. The energy preservation of the inviscid form is no longer true in the

context of (4.1) as energy will dissipate due to the viscous terms. The precise rate of energy dissipation is given by the one-dimensional form of the estimate (4.10). One of the key aims of designing energy preserving schemes is to obtain a faithful discretization of this energy dissipation balance.

We start by writing down the following scheme for the one-dimensional shallow water equations with eddy viscosity in semi-discrete form,

$$\frac{dU_i(t)}{dt} + \frac{1}{\Delta x} \left(\tilde{F}_{i+\frac{1}{2}} - \tilde{F}_{i-\frac{1}{2}} \right) = \frac{\nu}{\Delta x} \left(Q_{i+\frac{1}{2}} - Q_{i-\frac{1}{2}} \right). \quad (4.28)$$

Here, $F_{i+\frac{1}{2}}$ is any energy preserving flux so that (4.18) holds, and $Q_{i+\frac{1}{2}}$ is a discretization of the viscous terms:

$$Q_{i+\frac{1}{2}} = \left[0, \bar{h}_{i+\frac{1}{2}} \left(\frac{u_{i+1} - u_i}{\Delta x} \right) \right]^\top. \quad (4.29)$$

Note that (4.28) is a consistent discretization of (4.1) in one space dimension.

Lemma 4.1 *Consider the viscous form of the shallow water equations (4.1) in one space dimension. Let $E := \frac{1}{2} (hu^2 + gh^2)$ denote the energy and let $\tilde{F}_{i+\frac{1}{2}}$ be any consistent, energy preserving numerical flux satisfying (4.18). Then, the solution U_i of the “eddy viscosity” scheme (4.28),(4.29) satisfies the discrete energy dissipation*

$$\begin{aligned} \frac{d}{dt} E(U_i(t)) = & -\frac{1}{\Delta x} \left(\tilde{H}_{i+\frac{1}{2}} - \tilde{H}_{i-\frac{1}{2}} \right) \\ & - \frac{\nu}{2} \left(\bar{h}_{i+\frac{1}{2}} \left(\frac{u_{i+1} - u_i}{\Delta x} \right)^2 + \bar{h}_{i-\frac{1}{2}} \left(\frac{u_i - u_{i-1}}{\Delta x} \right)^2 \right), \end{aligned} \quad (4.30)$$

where

$$\tilde{H}_{i+\frac{1}{2}} := \left\langle \bar{V}_{i+\frac{1}{2}}, F_{i+\frac{1}{2}} \right\rangle - \bar{\Psi}_{i+\frac{1}{2}} - \bar{h}_{i+\frac{1}{2}} \bar{u}_{i+\frac{1}{2}} \frac{\llbracket u_{i+\frac{1}{2}} \rrbracket}{\Delta x}.$$

Summing over all i , we get the following energy dissipation estimate:

$$\frac{d}{dt} \sum_i E_i \equiv -\frac{\nu}{2} \sum_i (h_i + h_{i+1}) \left(\frac{u_{i+1} - u_i}{\Delta x} \right)^2.$$

Hence the finite volume scheme (4.28) leads to a discrete form of the exact energy balance (4.10).

The proof follows by direct calculations with energy preserving fluxes; see Theorem 5.1 in Tadmor and Zhong (2008) for details.

Thus, any energy preserving flux together with a central discretization of the viscous terms as in (4.28) results in a faithful discretization of the energy balance for the viscous form of the shallow water equations. We will check this fact and study its implications by considering the following numerical experiment.

Numerical experiment with eddy viscosity

We consider the viscous form (4.1) in one space dimension together with the viscosity parameter $\nu = 0.01$ and the initial data given by the one-dimensional dam-break problem (4.27). Since the results with all the three energy preserving fluxes were so similar, we only report on the results obtained with the EEC scheme, together with a central discretization of viscous terms (4.28). Time integration is performed with an RK2 scheme with a CFL number that takes into account the viscous terms in the scheme. We choose the CFL number of 0.1 in subsequent computations. The results in Figure 4.4 show that the EEC scheme to-

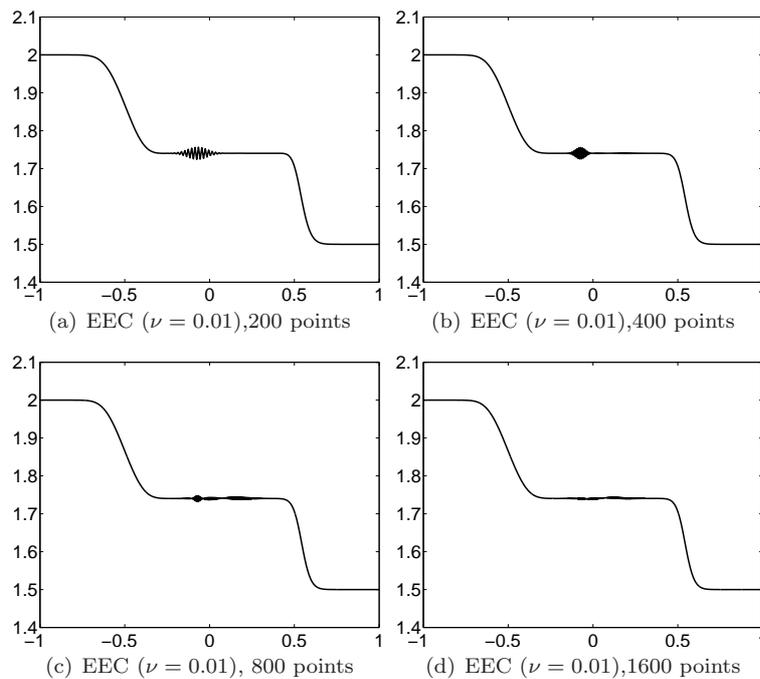


Fig. 4.4. Height h at time $t = 0.4$ computed with EEC scheme, RK2 at CFL 0.1 and central discretization of viscous terms (4.28) with eddy viscosity $\nu = 0.01$ on four different meshes.

gether with eddy viscosity results in a stable and robust approximation of the viscous shallow water equations. The right-going shock is dissipated as we are using a relatively large eddy viscosity $\nu = 0.01$. These results should be contrasted with plots from inviscid schemes shown in Figure 4.3. We see that the addition of the physical eddy viscosity leads to a dramatic reduction of the oscillations generated by the energy preserving scheme. Clearly, the addition of physical viscosity leads to an energy dissipation of the form (4.30), and the energy balance is faithfully discretized. However, we see from Figure 4.4 that small amplitude, high-frequency oscillations are still present on under-resolved meshes. As the mesh is refined, the amplitude of these oscillations is further reduced and the oscillations are hardly visible on a fine mesh of 1600 points. This behavior is consistent with expectations: the scales in the computed solution are given in terms of a *mesh Reynolds number* and we need to resolve them in order to obtain a solution without any oscillations. The mesh Reynolds number in this example appears to yield a mesh of roughly 1600 points to resolve all the scales in the viscous problem. An extreme illustration of the effect of eddy viscosity is obtained by comparing the approximate heights computed on 1600 mesh points with the EEC scheme and $\nu = 0$ (inviscid) and $\nu = 0.01$ (viscous) eddy viscosities. We consider the bottom right figures in Figure 4.3 and 4.4, respectively, and observe that the addition of viscosity dampens the oscillations dramatically.

4.2.7 Numerical Diffusion: Numerical Experiments 4–5

The previous section clearly illustrates the role of eddy viscosity in designing a stable numerical scheme for shallow water equations. The key observation is the need to resolve the viscous scales of the problem. The use of physical eddy viscosity is constrained by the fact that eddy viscosity in real physical applications is expected to be very small (smaller than 10^{-2}). This implies that we need to compute on very fine meshes in order to resolve all the viscous scales. Since this necessitates using large computational resources, we need to find an alternative way of designing stable numerical schemes for real world applications.

The standard way of designing stable schemes for problems with shocks in the finite volume framework is to add numerical diffusion. The numerical diffusion is built into the structure of numerical fluxes. In fact, standard finite volume fluxes $\tilde{F}_{i+\frac{1}{2}}$ for the system (4.17) based on 3-point

stencils can be written in the viscous form (consult Tadmor (1984))

$$F_{i+\frac{1}{2}} = F(U_i, U_{i+1}) = \frac{1}{2}(F(U_i) + F(U_{i+1})) - Q_{i+\frac{1}{2}}(V_{i+1} - V_i), \quad (4.31)$$

where $Q_{i+\frac{1}{2}}$ is a suitable numerical diffusion matrix coefficient. One of the fundamental results of Tadmor (1987) was to provide a criterion for whether the flux in (4.31) is entropy stable or not. We restate this result in the following lemma.

Lemma 4.2 (Tadmor (1987)) *Assume that the system of conservation laws (4.14) is equipped with an entropy/entropy flux pair (E, F) . Let $F_{i+\frac{1}{2}}$ be a finite volume numerical flux consistent with (4.14), and let $\tilde{F}_{i+\frac{1}{2}}$ be an entropy conservative numerical flux so that (4.18) holds. Let $Q_{i+\frac{1}{2}}$ and $\tilde{Q}_{i+\frac{1}{2}}$ be the corresponding numerical diffusions associated with $F_{i+\frac{1}{2}}$ and $\tilde{F}_{i+\frac{1}{2}}$, respectively. If*

$$Q_{i+\frac{1}{2}} \geq \tilde{Q}_{i+\frac{1}{2}} \quad \forall i \quad (4.32)$$

(in the sense that $Q_{i+\frac{1}{2}} - \tilde{Q}_{i+\frac{1}{2}}$ is a symmetric positive definite matrix), then the scheme

$$\frac{dU_i(t)}{dt} = -\frac{1}{\Delta x} (F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}) \quad (4.33)$$

is entropy stable, i.e., $\frac{d}{dt} \sum_i E(U_i(t)) \leq 0$.

In words — a finite volume flux, with *more* numerical viscosity than the entropy conservative flux, necessarily dissipates entropy and hence is stable. Several examples of entropy preserving fluxes were specified in Tadmor (1987) and in Tadmor (2003). We shall mention two in the context of the shallow water system (4.13).

(i) *Rusanov flux*. This energy stable flux takes the form,

$$F_{i+\frac{1}{2}}^{Rus} = \frac{1}{2}(F(U_i) + F(U_{i+1})) - \max_k \{ |(\lambda_k)_i|, |(\lambda_k)_{i+1}| \} (U_{i+1} - U_i), \quad (4.34)$$

where $(\lambda_k)_i$ are the eigenvalues of $F'(U_i)$,

$$(\lambda_1)_i = u_i - \sqrt{gh_i}, \quad (\lambda_2)_i = u_i + \sqrt{gh_i}.$$

Note that the diffusion is scaled by a local estimate of the speed of propagation.

(ii) *Roe flux.* This well-known flux is given by

$$F_{i+\frac{1}{2}}^{Roe1} = \frac{1}{2}(F(U_i) + F(U_{i+1})) - R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^{-1} (U_{i+1} - U_i), \quad (4.35)$$

where

$$|\Lambda_{i+\frac{1}{2}}| = \text{diag}\{ |(\lambda_1)_{i+\frac{1}{2}}|, |(\lambda_2)_{i+\frac{1}{2}}| \}$$

and

$$R_{i+\frac{1}{2}} = \begin{bmatrix} 1 & 1 \\ (\lambda_1)_{i+\frac{1}{2}} & (\lambda_2)_{i+\frac{1}{2}} \end{bmatrix}$$

for some suitable average states $U_{i+\frac{1}{2}}$. It was shown in Tadmor (1987) that the Roe flux, together with some appropriate entropy fix, is entropy stable. The choice of the parameters in the entropy fix was described in detail in Tadmor (2003).

It is well-known that the Rusanov flux (4.34) can be too diffusive for practical applications. The Roe flux (4.35) is much less diffusive and the shocks are captured sharply. In particular, choosing

$$h_{i+\frac{1}{2}} = \frac{h_i + h_{i+1}}{2}, \quad u_{i+\frac{1}{2}} = \frac{\sqrt{h_i}u_i + \sqrt{h_{i+1}}u_{i+1}}{\sqrt{h_i} + \sqrt{h_{i+1}}},$$

as the average states in (4.35) ensures that isolated single shocks are resolved exactly. However, the Roe flux lacks entropy stability and we refer to Tadmor (2003) for a discussion on entropy stable modifications of the Roe flux.

Here we propose a novel choice of a numerical diffusion flux for (4.13), in the spirit of a Roe flux, which is energy stable. The resulting numerical flux is designed in two steps. In the first step, we will replace the central part of the flux, $(f(U_i) + f(U_{i+1}))/2$, with the energy preserving fluxes satisfying (4.18); we can choose any of the three energy preserving fluxes for the shallow water equation. The next step is to design a numerical diffusion operator. Here, we need the following simple lemma.

Lemma 4.3 *Consider the shallow water equations in one dimension, together with energy $E = (gh^2 + hu^2)/2$ and the energy variables $V = [gh - \frac{u^2}{2}, u]^T$. Let the values U_i, U_{i+1} across an interface be given.*

(i) *We have the identity*

$$[[U_{i+\frac{1}{2}}]] = (\overline{U_V})_{i+\frac{1}{2}} [[V_{i+\frac{1}{2}}]], \quad (4.36)$$

where

$$(\overline{U_V})_{i+\frac{1}{2}} = \frac{1}{g} \begin{bmatrix} 1 & \overline{u}_{i+\frac{1}{2}} \\ \overline{u}_{i+\frac{1}{2}} & (\overline{u^2})_{i+\frac{1}{2}} + g\overline{h}_{i+\frac{1}{2}} \end{bmatrix}$$

and $\overline{h}_{i+\frac{1}{2}}$ and $\overline{u}_{i+\frac{1}{2}}$ are the arithmetic averages across the interface.

- (ii) Define the scaled matrix of right eigenvectors of $F'(U_{i+\frac{1}{2}})$ for some averaged state $U_{i+\frac{1}{2}}$:

$$R_{i+\frac{1}{2}} = \frac{1}{\sqrt{2g}} \begin{bmatrix} 1 & 1 \\ (\lambda_1)_{i+\frac{1}{2}} & (\lambda_2)_{i+\frac{1}{2}} \end{bmatrix}. \quad (4.37)$$

Then we have

$$R_{i+\frac{1}{2}} R_{i+\frac{1}{2}}^\top = (U_V)_{i+\frac{1}{2}}, \quad (4.38)$$

where U_V is the symmetric positive definite change of variables matrix

$$(U_V)_{i+\frac{1}{2}} = \frac{1}{g} \begin{bmatrix} 1 & u_{i+\frac{1}{2}} \\ u_{i+\frac{1}{2}} & u_{i+\frac{1}{2}}^2 - gh_{i+\frac{1}{2}} \end{bmatrix}$$

evaluated at the same average state.

The proof of the above identities follows by a straightforward calculation which we omit. Note that part (i) of the lemma above provides an appropriate average state at which the jump of the conservative variables can be expressed in terms of the jump of the energy variables. This state is not the arithmetic average of the conservative variables, but rather of the primitive variables. Part (ii) provides a suitable scaling for the eigenvectors.

This lemma can be used to design a suitable Roe-type diffusion operator. We consider the usual Roe diffusion, $Q_{i+\frac{1}{2}}^{Roe1}$ in (4.35), evaluated at the average states in (4.36) and use Lemma 4.3 to obtain,

$$\begin{aligned} Q_{i+\frac{1}{2}}^{Roe1} [U_{i+\frac{1}{2}}] &= R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^{-1} [U_{i+\frac{1}{2}}], \\ &= R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^{-1} (\overline{U_V})_{i+\frac{1}{2}} [V_{i+\frac{1}{2}}], \quad (\text{by (4.36)}) \\ &= R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^{-1} R_{i+\frac{1}{2}} R_{i+\frac{1}{2}}^\top [V_{i+\frac{1}{2}}], \quad (\text{by (4.38)}) \\ &= R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^\top [V_{i+\frac{1}{2}}]. \end{aligned}$$

These formal calculations suggest that we can choose a numerical diffusion operator in terms of the energy variables with proper scaling of

the eigenvectors and evaluation of the eigenvalues and eigenvectors at an appropriate state.

Now, let $\tilde{F}_{i+\frac{1}{2}}$ be any energy preserving numerical flux (4.18). We introduce the following Roe-type flux:

$$F_{i+\frac{1}{2}}^{ERoe1} = F^{ERoe1}(U_i, U_{i+1}) := \tilde{F}_{i+\frac{1}{2}} - \frac{1}{2} R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^\top \llbracket V_{i+\frac{1}{2}} \rrbracket, \quad (4.39)$$

where $R_{i+\frac{1}{2}}$ is as defined in (4.37), and

$$|\Lambda_{i+\frac{1}{2}}| = \text{diag} \left\{ |\bar{u}_{i+\frac{1}{2}} - \sqrt{g\bar{h}_{i+\frac{1}{2}}}|, |\bar{u}_{i+\frac{1}{2}} + \sqrt{g\bar{h}_{i+\frac{1}{2}}}| \right\}.$$

$F_{i+\frac{1}{2}}^{ERoe1}$ is clearly consistent. It differs from the usual Roe flux (4.35) in that the central term is an energy preserving flux, and that the diffusion is given in terms of energy variables rather than conservative variables. We refer to $F_{i+\frac{1}{2}}^{ERoe1}$ as *entropic Roe flux*: indeed, the associated scheme is entropy stable as quantified in the following theorem.

Theorem 4.2 *Let U_i be the solution of the entropic Roe scheme*

$$\frac{dU_i(t)}{dt} = -\frac{1}{\Delta x} \left(F_{i+\frac{1}{2}}^{ERoe1} - F_{i-\frac{1}{2}}^{ERoe1} \right). \quad (4.40)$$

Then the following discrete energy estimate holds:

$$\begin{aligned} \frac{d}{dt} (E(U_i)(t)) &= -\frac{1}{\Delta x} (\tilde{H}_{i+\frac{1}{2}} - \tilde{H}_{i-\frac{1}{2}}) \\ &\quad - \frac{1}{4\Delta x} \left\langle \llbracket V_{i+\frac{1}{2}} \rrbracket, R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^\top \llbracket V_{i+\frac{1}{2}} \rrbracket \right\rangle \\ &\quad - \frac{1}{4\Delta x} \left\langle \llbracket V_{i-\frac{1}{2}} \rrbracket, R_{i-\frac{1}{2}} |\Lambda_{i-\frac{1}{2}}| R_{i-\frac{1}{2}}^\top \llbracket V_{i-\frac{1}{2}} \rrbracket \right\rangle, \end{aligned} \quad (4.41)$$

where

$$\tilde{H}_{i+\frac{1}{2}} := \left\langle \bar{V}_{i+\frac{1}{2}}, \tilde{F}_{i+\frac{1}{2}} \right\rangle - \bar{\Psi}_{i+\frac{1}{2}} + \frac{1}{2} \left\langle \bar{V}_{i+\frac{1}{2}}, R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^\top \llbracket V_{i+\frac{1}{2}} \rrbracket \right\rangle.$$

Summing over all i , we get the following energy dissipation estimate:

$$\frac{d}{dt} \sum_i E_i = -\frac{1}{2\Delta x} \sum_i \left\langle \llbracket V_{i+\frac{1}{2}} \rrbracket, R_{i+\frac{1}{2}} |\Lambda_{i+\frac{1}{2}}| R_{i+\frac{1}{2}}^\top \llbracket V_{i+\frac{1}{2}} \rrbracket \right\rangle.$$

In particular, since the matrix $R|\Lambda|R^\top$ is symmetric non-negative definite, the scheme (4.40) is energy stable.

The proof of this theorem is a straightforward generalization of the proof of entropy stability with numerical diffusion in entropy variables (consult Tadmor (1987),(2003)): one multiplies both sides of (4.33) by V_i and

summation by parts while taking into account the energy preserving flux and the special form of the diffusion matrix in (4.39) yields the energy estimate (4.41).

Remark 4.2 The idea of defining the numerical diffusion operator for finite volume fluxes in terms of entropy variables is not new but was proposed in Tadmor (1986),(1987) and subsequently, was used in Hughes Franca and Mallet (1986), Khalfallah and Lerat (1989) and others. The specific structure of the diffusion operator in (4.39), however, is novel for the shallow water equations. The flux (4.39) is less dissipative than the Rusanov flux and does not need an entropy fix like the Roe flux.

Numerical experiments with numerical diffusion

We present a series of one-dimensional numerical experiments of schemes with added numerical diffusion. We will test both the standard Rusanov and Roe schemes (4.34) and (4.35), and compare the results with the new entropy stable Roe-type scheme (4.39). The latter employs an energy preserving flux $\tilde{F}_{i+\frac{1}{2}}$ and we chose to use here the EEC flux (4.26). The resulting entropic scheme (4.39), (4.26) is denoted *ERoe*. We start with the one-dimensional dam-break problem with initial data (4.27) and domain $[-1, 1]$. We compute the approximate solutions with the Roe, Rusanov and ERoe schemes on a uniform mesh with 100 mesh points and plot the approximate heights in Figure 4.5. As shown in this figure,

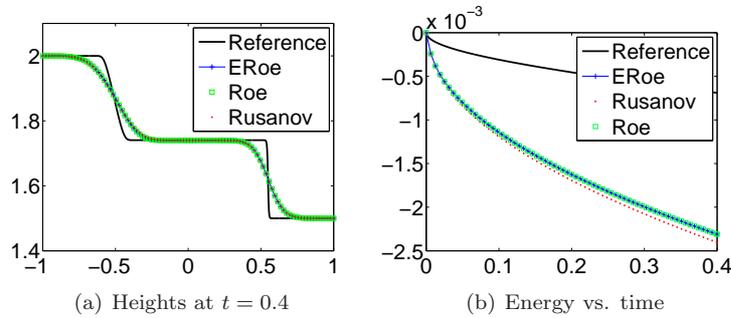


Fig. 4.5. Solutions computed with the Roe, Rusanov and ERoe schemes for 1D dam break problem with 100 mesh points. Left: Height at $t = 0.4$, Right: Energy vs. time.

the three schemes behave as expected. The Rusanov scheme is slightly more diffusive than the Roe-type schemes, with the right-going shock

being slightly more smeared. However, the difference is not much. Both the Roe and ERoe schemes compute the solution quite well at this coarse mesh resolution. The difference between the two schemes is negligible for this problem. Figure 4.5(b) shows the energy history, and we observe that both the Roe and ERoe schemes dissipate energy in an identical manner. The Rusanov scheme dissipates more energy than either of the two Roe-type schemes. All three schemes are energy-stable, in that they diffuse more energy than the physical solution.

The above numerical experiment illustrated that the new ERoe scheme is robust and very similar in behavior to the standard Roe scheme. The ERoe scheme, however, was proved to be energy stable, thus providing an entropy fix for the Roe scheme. The following two numerical experiments will illustrate these points.

Numerical experiment #4: A dam-break problem

We consider a different one-dimensional dam-break problem for the shallow water equations with initial data

$$h(x, 0) = \begin{cases} 15 & \text{if } x < 0, \\ 1 & \text{if } x > 0, \end{cases} \quad u(x, 0) \equiv 0 \quad (4.42)$$

in the computational domain $[-1, 1]$, and we set $g = 1$. The difference between the above initial data and the one in experiment #1 in (4.27) is the large initial jump in height. The exact solution still consists of a right going shock and a left going rarefaction, separated by a region of a constant height of 5. We compute the solutions with the Roe and ERoe schemes on a mesh of 100 points up to time $t = 0.15$. We also compute the Rusanov scheme and a reference solution on a finer mesh of 3200 points. As shown in Figure 4.6, the ERoe scheme provides a good approximation to the exact solution. The Roe scheme approximates the shock equally well; however it produces a spurious steady shock at approximately $x = 0$. The magnitude of this shock is about 2.5 and it is entirely unphysical. This behavior of spurious waves generated by the Roe scheme is well known (see LeVeque (2002)), and the scheme needs to be entropy fixed, as in, e.g., Tadmor (2003). This is a key difference between the Roe and ERoe schemes. They seem to have the same resolution, but the Roe scheme (at least in its non-entropy fixed version) is unstable in some cases.

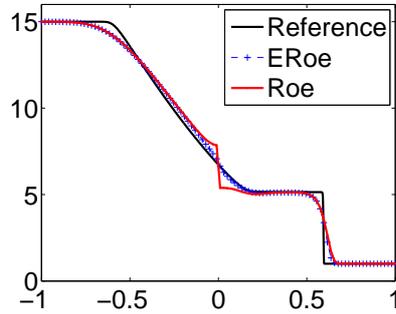


Fig. 4.6. Heights computed at time $t = 0.4$ with the Roe and ERoe schemes for numerical experiments with 100 mesh points.

Numerical experiment #5: An expansion problem

Another well documented issue with the standard Roe scheme is its lack of positivity, i.e., it may produce negative heights. The ERoe scheme might also suffer from this problem. We present an experiment where the Roe scheme leads to negative heights, whereas the ERoe scheme retains positivity. We consider initial data

$$h(x, 0) \equiv 1, \quad u(x, 0) = \begin{cases} -4 & \text{if } x < 0, \\ 4 & \text{if } x > 0. \end{cases} \quad (4.43)$$

The computational domain is $[-1, 1]$ with transparent boundary conditions and $g = 1$. The initial velocity is chosen such that the fluid is pushed away from the center of the domain in both directions, leading to the formation of an almost dry zone, with very low values of height in the center of the domain. We compute the solutions with the Roe and ERoe schemes for a uniform mesh of 100 mesh points. The results from Figure 4.7 show that the ERoe scheme approximates the solution quite well and retains positivity of the height in the near-dry zone around $x = 0$. However, the Roe scheme fails in this case due to negative heights which form at $t \sim 0.006$ (we therefore show the computed height just before this failure). This illustrates how the ERoe stabilizes the positivity failure of the Roe scheme; we do not, however, claim that the ERoe is positivity preserving, and one might expect it to fail if we increase the velocity in (4.43).

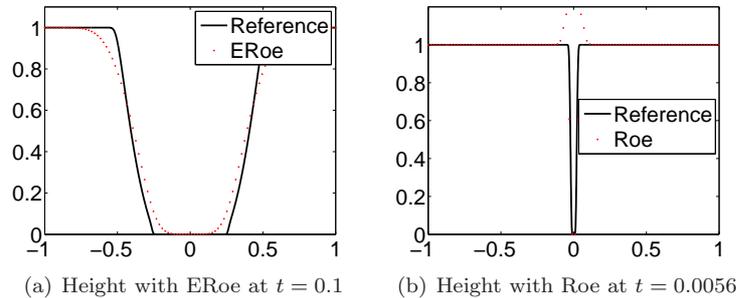


Fig. 4.7. Solutions computed with the Roe, and ERoe schemes for 1D expansion problem with 100 mesh points. Left: Height at $t = 0.4$ with ERoe, Right: Height at $t = 0.01$ with Roe.

Computational costs

The above examples illustrate that the ERoe is quite robust and energy stable. It is as accurate as the standard Roe scheme and is more stable with respect to energy and positivity. A natural question is whether this increased stability comes at a price. Hence, we consider the standard one-dimensional dam-break problem with initial data (4.27) and compute the solutions with Roe, ERoe and Rusanov schemes. We compute a reference solution with the Rusanov scheme on a fine mesh of 3200 mesh points and consider L^1 errors in height with respect to this reference solution. Next, we compute with three different mesh resolutions and find the meshes on which each of the schemes will yield a relative error of 1, 0.5 and 0.1 percent, respectively. The run times are then normalized such that unit time corresponds to 0.0205 seconds. We present the normalized run times with all the three schemes in Table 4.2. We see

Relative Energy error	1	0.5	0.1
Rusanov	1.05	8.24	203.41
Roe	1.15	8.43	208.29
ERoe	1	7.36	171.7

Table 4.2. Normalized run-times for the Rusanov, Roe and ERoe schemes on the one-dimensional dam-break problem with three different levels of relative error in height.

from the table that the ERoe scheme has the lowest computational cost for the same level of error in all the three different error levels. Surpris-

ingly the Rusanov scheme is less costly and hence more efficient than the Roe scheme. The Roe scheme is about 15 to 20 percent more expensive in this example than the ERoe scheme. This might be attributed to the fact that the diffusion operator in the ERoe scheme might involve less computational work than the diffusion operator in the Roe scheme. Hence, all the above tests indicate that our proposed ERoe scheme is actually as accurate, more stable and less computationally expensive than the standard Roe scheme.

4.2.8 Second-order Extensions: Numerical Experiment 6

The ERoe flux (4.39) has two parts: A second-order symmetric energy preserving part and a numerical diffusion operator. The numerical diffusion is first-order accurate in space and hence, the overall accuracy of the ERoe scheme is restricted to first-order. This leads to smeared shocks, as we observed in the previous numerical experiments. It is straightforward to extend this scheme to obtain second-order accuracy in space. For this purpose, we follow the fairly standard approach and replace the piecewise-constant cell averages U_i in (4.33) with a non-oscillatory piecewise linear reconstruction, e.g., Kurganov and Tadmor (2002), of the form

$$p_i(x) = U_i + \frac{U'_i}{\Delta x}(x - x_i). \quad (4.44)$$

The numerical derivative U'_i is given by

$$U'_i := \text{minmod} \left(U_{i+1} - U_i, \frac{1}{2}(U_{i+1} - U_{i-1}), U_i - U_{i-1} \right), \quad (4.45)$$

where minmod is the function

$$\text{minmod}(a, b, c) := \begin{cases} \text{sgn}(a) \min\{|a|, |b|, |c|\}, & \text{if } \text{sgn}(a) = \text{sgn}(b) = \text{sgn}(c) \\ 0, & \text{otherwise.} \end{cases}$$

The reconstruction is fairly standard, and we can use other limiters for the numerical derivatives. Let

$$U_i^E = p_i(x_{i+\frac{1}{2}}), \quad U_i^W = p_i(x_{i-\frac{1}{2}}).$$

The second-order version of the ERoe flux (4.39) is then

$$F_{i+\frac{1}{2}}^{ERoe2} = \tilde{F}(U_i^E, U_{i+1}^W) - \frac{1}{2}R|\Lambda|R^\top(V_{i+1}^W - V_i^E), \quad (4.46)$$

where \tilde{F} is any energy preserving flux (4.18), V_i^E, V_i^W are the energy variables evaluated at U_i^E and U_i^W , and the matrices R and Λ are the

corresponding matrices in (4.36) and (4.38) with respect to the reconstructed values U_i^E and U_{i+1}^W . The resulting finite volume scheme with this flux is formally second-accurate in space. However, we can no longer prove that it is energy stable like the first-order version. Instead we will test with this second-order flux in our next numerical experiment.

Numerical experiment #6: Second-order computation of dam-break problem

We consider the standard one-dimensional dam-break problem with initial data (4.27) and consider the first-order (4.39) and second-order (4.46) versions of the ERoe scheme. We compute heights on a uniform mesh of 100 mesh points. Both schemes are integrated in time using a standard second-order RK2 method with a CFL number of 0.45, and we show the computed heights at time $t = 0.4$ in Figure 4.8. This figure

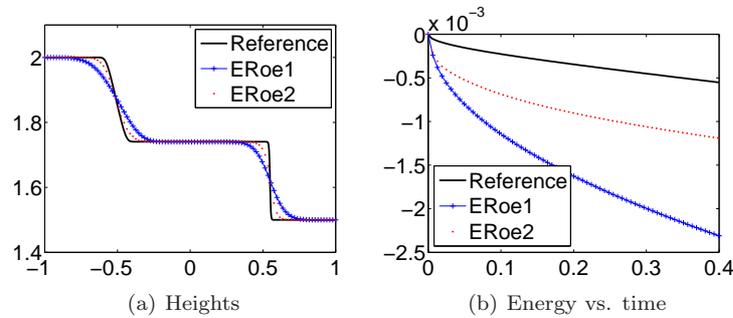


Fig. 4.8. Solutions computed with the first and second-order versions of the ERoe scheme with 100 mesh points.

clearly shows the effect of a higher order of accuracy. The shock and rarefaction are both resolved to a much better extent with the second-order scheme. Furthermore, the energy dissipation with the second-order scheme is much less than that of the first-order scheme. The second-order scheme seems to be energy stable in this example although we were unable to obtain a proof of this. Even higher order extensions can be performed by using fairly standard ENO and WENO type reconstructions.

4.3 The Two-Dimensional Problem

The analysis and numerics presented in Section 4.2 for the one-dimensional problem can be easily extended to the two-dimensional shallow water equations (4.2). We will be brief in the exposition, as most of the details are similar to the one-dimensional case. To begin with, the energy variables for the two-dimensional form of the equations (4.2) and the associated potentials are given by

$$V = \left[gh - \frac{u^2 + v^2}{2}, u, v \right]^\top, \quad \Psi = \frac{1}{2}guh^2, \quad \Phi = \frac{1}{2}gvh^2, \quad (4.47)$$

where Ψ and Φ are the energy potentials in the x - and y -directions, respectively.

4.3.1 Energy Preserving Schemes: Numerical Experiment 7

We seek to approximate the two-dimensional shallow water system (4.2) with a standard finite volume scheme of the form (4.12). As in the one-dimensional case, the energy preserving (entropy preserving for a general system (4.3)) finite volume fluxes are characterized in Tadmor (1987), Tadmor (2003), as fluxes satisfying the following conditions:

$$\left\langle [V_{i+\frac{1}{2},j}], \tilde{F}_{i+\frac{1}{2},j} \right\rangle = [\Psi_{i+\frac{1}{2},j}], \quad \left\langle [V_{i,j+\frac{1}{2}}], \tilde{G}_{i,j+\frac{1}{2}} \right\rangle = [\Phi_{i,j+\frac{1}{2}}]. \quad (4.48)$$

Consistent numerical fluxes F, G that satisfy (4.48) are energy preserving for the shallow water equations, and we obtain a two-dimensional version of Theorem 4.1 (we skip the details and refer the reader to Tadmor (1987) for details).

In the one-dimensional case, we obtained three different schemes satisfying (4.18). The flux of the form (4.20) can be similarly defined for the two-dimensional case. However, we failed to explicitly compute the averages in phase space and obtain a simple explicit flux as we did in the one-dimensional case. This is largely due to the increased complexity of the resulting algebraic expressions. Hence, we do not have a flux analogous to the AEC flux (4.21) for the two-dimensional case.

The pathwise approach of Tadmor (2003) can be easily extended to cover the two-dimensional case, and has been done so in Tadmor and Zhong (2008). We skip the details of the flux and refer the reader to Algorithm 3.1 in Tadmor and Zhong (2008) for the description of this flux. The pathwise flux is a natural generalization of the PEC flux (4.23), and we continue referring to the resulting scheme as the PEC scheme.

It is straightforward to extend the explicit differencing based EEC flux (4.26) to the two-dimensional case. We carry out the steps indicated in the derivation of (4.26) for the two-dimensional case, and obtain the following entropy conservative flux:

$$\tilde{F}_{i+\frac{1}{2},j} = \begin{bmatrix} \bar{h}_{i+\frac{1}{2},j} \bar{u}_{i+\frac{1}{2},j} \\ \bar{h}_{i+\frac{1}{2},j} (\bar{u}_{i+\frac{1}{2},j})^2 + \frac{g}{2} (\bar{h}^2)_{i+\frac{1}{2},j} \\ \bar{h}_{i+\frac{1}{2},j} \bar{u}_{i+\frac{1}{2},j} \bar{v}_{i+\frac{1}{2},j} \end{bmatrix}, \quad (4.49a)$$

and

$$\tilde{G}_{i,j+\frac{1}{2}} = \begin{bmatrix} \bar{h}_{i,j+\frac{1}{2}} \bar{v}_{i,j+\frac{1}{2}} \\ \bar{h}_{i,j+\frac{1}{2}} \bar{u}_{i,j+\frac{1}{2}} \bar{v}_{i,j+\frac{1}{2}} \\ \bar{h}_{i,j+\frac{1}{2}} (\bar{v}_{i,j+\frac{1}{2}})^2 + \frac{g}{2} (\bar{h}^2)_{i,j+\frac{1}{2}} \end{bmatrix}. \quad (4.49b)$$

It is easy to check that the above fluxes satisfy (4.48) and are consistent. This flux is also trivial to implement in code as we simply need to evaluate averages. It is symmetric in its arguments and reduces to (4.26) for the one-dimensional case. We will denote the resulting scheme with (4.49) as the EEC scheme.

Both the PEC scheme and EEC scheme are energy preserving for the two-dimensional shallow water equations. We compare their numerical behavior in the following numerical experiment.

Numerical Experiment #7: 2D cylindrical dam-break

We consider the shallow water equations (4.2) in the domain $[-1, 1] \times [-1, 1]$ with initial data

$$h(x, y, 0) = \begin{cases} 2, & \text{if } \sqrt{x^2 + y^2} < 0.5 \\ 1, & \text{otherwise} \end{cases}, \quad u(x, y, 0) \equiv 0. \quad (4.50)$$

The initial data represents the breaking of a cylindrical dam, and the problem has radial symmetry. We set $g = 1$. The exact solution consists of a circular shock moving out and a rarefaction moving inward. We compute with both EEC and PEC schemes on a uniform 50×50 mesh. The computations are performed with an RK2 time integration routine and a CFL number of 0.45.

As shown in Figure 4.9, the energy preserving schemes show very little difference in the numerical results. As in the one-dimensional cases, there are oscillations for both schemes as the energy is redistributed into smaller scales, and dispersive effects dominate due to the lack of any diffusive mechanism. We also show the energy errors in time in Figure 4.10. The results from Figure 4.10 show that the energy errors for

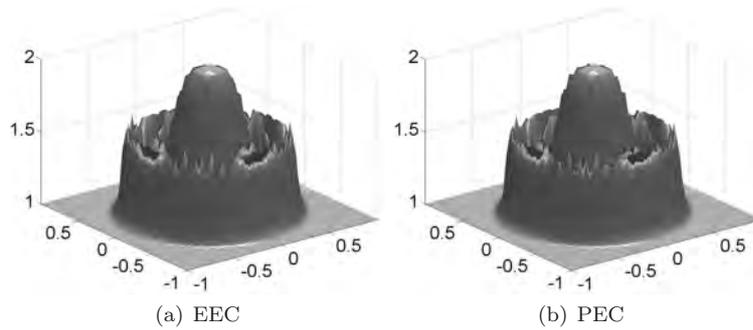


Fig. 4.9. Approximate heights for the cylindrical dam-break problem at time $t = 0.2$ computed on a uniform 50×50 mesh with both EEC and PEC schemes.

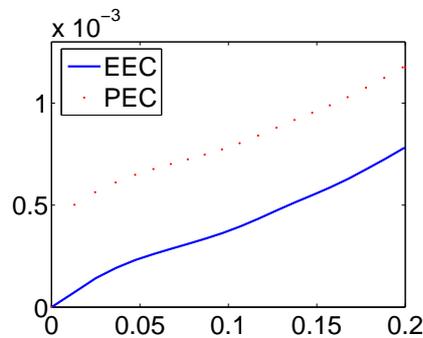


Fig. 4.10. Energy vs. time for the 2-d cylindrical dam-break problem and EEC, PEC schemes for a 50×50 mesh and RK2 with CFL 0.45.

both schemes are quite low for this extremely coarse mesh and second-order time integration with a moderately high CFL number. The results are consistent with those observed for the one-dimensional problem in Section 4.2. However, it seems that the PEC scheme generates a larger error in energy of about 2-3 times more than the EEC scheme.

The errors in energy are purely due to the time integration. As in one dimension, we observe a considerable reduction of the energy error by decreasing the time step. A sample computation is presented in Figure 4.11, where the energy errors generated with the EEC scheme and an RK2 method for two different CFL numbers is shown. The results in Figure 4.11 show that halving the CFL number (i.e., halving the time step) reduces the energy error by a factor of eight. Thus, the energy errors behave like $\mathcal{O}(\Delta t)^3$, as in the one-dimensional case. As in the one-

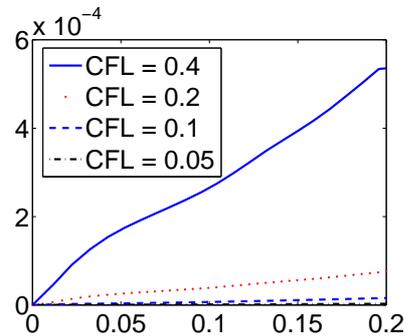


Fig. 4.11. Energy vs. time for the 2-d cylindrical dam-break problem and EEC scheme for a 100×100 mesh and RK2 with four different CFL numbers.

dimensional case, going to an RK3 time integration will further reduce the energy decay.

As stated before, energy preserving schemes lead to oscillations near shocks as there is no dissipative mechanism. In the one-dimensional case, these oscillations increased in frequency forming a modulated behavior as the mesh was refined. The same effects hold true in the two-dimensional case. We illustrate this behavior by presenting the approximate heights computed with the EEC scheme for a 100×100 mesh and a 200×200 mesh with RK2 time-integration in Figure 4.12. The figure clearly shows that increasing the mesh resolution results in oscillations of higher frequency although the magnitude of oscillations remains bounded. This is clear evidence of the dispersive behavior of the schemes in the absence of diffusion.

Computational costs

There are minor differences in the quality of the solutions obtained with the EEC and PEC schemes in two dimensions. The main difference lies in the simplicity of the EEC scheme and its low computational cost. As in one dimension, we illustrate the efficiency of the EEC scheme vis-à-vis the PEC scheme by fixing a uniform 50×50 mesh and lowering the time step to obtain energy errors of 10^{-3} , 10^{-4} and 10^{-5} , respectively. The resulting run times are then shown in Table 4.3. All the times are normalized so that unit time is taken to be 0.0617 seconds. As shown in Table 4.3, the PEC scheme is about 5 to 6 times more expensive than the EEC scheme. This is to be expected, as the PEC scheme was about 2.5 to 3 times more expensive in one dimension (see Table 4.1). The EEC

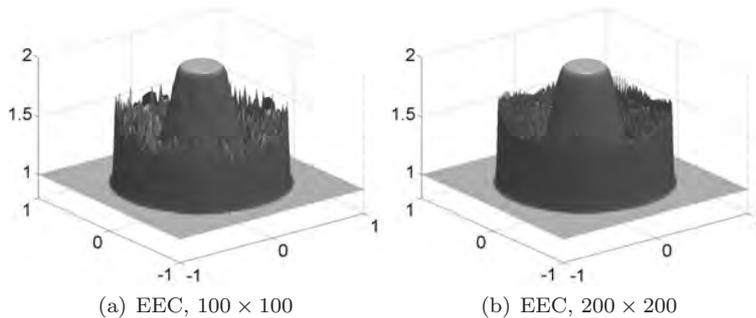


Fig. 4.12. Approximate heights for the cylindrical dam-break problem at time $t = 0.4$ computed on a uniform 100×100 and 200×200 meshes with the EEC and PEC scheme.

Energy error	10^{-3}	10^{-4}	10^{-5}
EEC	1	2.18	5.07
PEC	4.79	11.5	30.47

Table 4.3. *Normalized run-times for the two energy preserving schemes on the two-dimensional cylindrical dam-break problem with three different levels of error in energy.*

scheme provides a considerable speed up if one is interested in higher dimensional computations.

4.3.2 Eddy Viscosity: Numerical Experiment #8

By a simple generalization of (4.29), the energy preserving schemes can be used together with a central discretization of the viscous terms to obtain a scheme that is energy stable and with energy dissipating at a rate dictated by the viscous terms in (4.1). Without going into details, we observe that we will obtain a discrete form of (4.10) (a two-dimensional generalization of (4.30)). Thus, combining energy preserving schemes with a central discretization yields a faithful discretization of the continuous energy decay estimate.

We test this contention on a numerical example by considering the two-dimensional cylindrical dam-break problem (4.50) with eddy viscosity $\nu = 0.01$ and show the results in Figure 4.13. As expected, the presence of eddy viscosity serves to dramatically reduce the oscillations in the energy preserving schemes. Since the resulting solution has very low

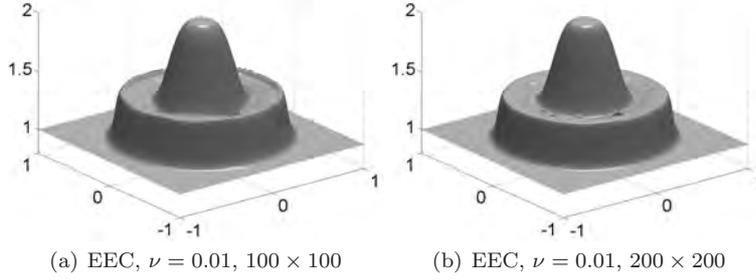


Fig. 4.13. Approximate heights for the cylindrical dam-break problem with eddy viscosity $\nu = 0.01$ at time $t = 0.4$ computed on uniform 100×100 and 200×200 meshes with the EEC scheme.

amplitude oscillations, it appears that the reasonably coarse 200×200 mesh is enough to resolve all the viscous scales. This is similar to what we observed in one dimension. Using an energy preserving scheme together with central viscous terms leads to an energy stable discretizations of the system, and the oscillations obtained for the inviscid problem are damped on a resolved mesh when one adds eddy viscosity.

4.3.3 Numerical Diffusion: Numerical Experiment #9

As in the one-dimensional case, we can use the energy preserving schemes to design suitable numerical diffusion operators leading to energy stable discretizations of the inviscid shallow water system (4.2), or the under-resolved viscous form (4.1). We imitate the strategy used in designing the energy stable Roe-type flux (4.39) in Section 4.2 and present its two-dimensional generalization. We start with the following lemma, which is a generalization of Lemma 4.3.

Lemma 4.4 *Consider the shallow water equations in two dimensions (4.2). Let the values $U_{i,j}, U_{i+1,j}, U_{i,j+1}$ across an interface be given.*

[i] *We have the identities*

$$\left[U_{i+\frac{1}{2},j} \right] = \overline{(U_V)}_{i+\frac{1}{2},j} \left[V_{i+\frac{1}{2},j} \right], \quad \left[U_{i,j+\frac{1}{2}} \right] = \overline{(U_V)}_{i,j+\frac{1}{2}} \left[V_{i,j+\frac{1}{2}} \right], \quad (4.51)$$

where

$$\overline{(UV)}_{i+\frac{1}{2},j} := \frac{1}{g} \begin{bmatrix} 1 & \bar{u}_{i+\frac{1}{2},j} & \bar{v}_{i+\frac{1}{2},j} \\ \bar{u}_{i+\frac{1}{2},j} & \bar{u}_{i+\frac{1}{2},j}^2 + g\bar{h}_{i+\frac{1}{2},j} & \bar{u}_{i+\frac{1}{2},j} \bar{v}_{i+\frac{1}{2},j} \\ \bar{v}_{i+\frac{1}{2},j} & \bar{u}_{i+\frac{1}{2},j} \bar{v}_{i+\frac{1}{2},j} & \bar{v}_{i+\frac{1}{2},j}^2 + g\bar{h}_{i+\frac{1}{2},j} \end{bmatrix}$$

and

$$\overline{(UV)}_{i,j+\frac{1}{2}} := \frac{1}{g} \begin{bmatrix} 1 & \bar{u}_{i,j+\frac{1}{2}} & \bar{v}_{i,j+\frac{1}{2}} \\ \bar{u}_{i,j+\frac{1}{2}} & \bar{u}_{i,j+\frac{1}{2}}^2 + g\bar{h}_{i,j+\frac{1}{2}} & \bar{u}_{i,j+\frac{1}{2}} \bar{v}_{i,j+\frac{1}{2}} \\ \bar{v}_{i,j+\frac{1}{2}} & \bar{u}_{i,j+\frac{1}{2}} \bar{v}_{i,j+\frac{1}{2}} & \bar{v}_{i,j+\frac{1}{2}}^2 + g\bar{h}_{i,j+\frac{1}{2}} \end{bmatrix}.$$

[ii] Define the scaled matrices of right eigenvectors of $F'(U_{i+\frac{1}{2},j})$ and $G'(U_{i,j+\frac{1}{2}})$ for averaged states $U_{i+\frac{1}{2},j}$ and $U_{i,j+\frac{1}{2}}$:

$$\begin{aligned} R_{i+\frac{1}{2},j}^x &= \frac{1}{\sqrt{2g}} \begin{bmatrix} 1 & 0 & 1 \\ \bar{u}_{i+\frac{1}{2},j} - \sqrt{g\bar{h}_{i+\frac{1}{2},j}} & 0 & \bar{u}_{i+\frac{1}{2},j} + \sqrt{g\bar{h}_{i+\frac{1}{2},j}} \\ \bar{v}_{i+\frac{1}{2},j} & \sqrt{g\bar{h}_{i+\frac{1}{2},j}} & \bar{v}_{i+\frac{1}{2},j} \end{bmatrix}, \\ R_{i,j+\frac{1}{2}}^y &= \frac{1}{\sqrt{2g}} \begin{bmatrix} 1 & 0 & 1 \\ \bar{u}_{i,j+\frac{1}{2}} & -\sqrt{g\bar{h}_{i,j+\frac{1}{2}}} & \bar{u}_{i,j+\frac{1}{2}} \\ \bar{v}_{i,j+\frac{1}{2}} - \sqrt{g\bar{h}_{i,j+\frac{1}{2}}} & 0 & \bar{v}_{i,j+\frac{1}{2}} + \sqrt{g\bar{h}_{i,j+\frac{1}{2}}} \end{bmatrix}. \end{aligned} \quad (4.52)$$

Then we have

$$R_{i+\frac{1}{2},j}^x \left(R_{i+\frac{1}{2},j}^x \right)^\top = \overline{(UV)}_{i+\frac{1}{2},j}, \quad R_{i,j+\frac{1}{2}}^y \left(R_{i,j+\frac{1}{2}}^y \right)^\top = \overline{(UV)}_{i,j+\frac{1}{2}}. \quad (4.53)$$

The proof follows from direct calculations. Using this lemma, we follow Section 4.2.7 and define the following Roe-type flux:

$$\begin{aligned} F_{i+\frac{1}{2},j}^{ERoe} &:= \tilde{F}_{i+\frac{1}{2},j} - \frac{1}{2} R_{i+\frac{1}{2},j}^x |\Lambda_{i+\frac{1}{2},j}^x| \left(R_{i+\frac{1}{2},j}^x \right)^\top [V_{i+\frac{1}{2},j}], \\ G_{i,j+\frac{1}{2}}^{ERoe} &:= \tilde{G}_{i,j+\frac{1}{2}} - \frac{1}{2} R_{i,j+\frac{1}{2}}^y |\Lambda_{i,j+\frac{1}{2}}^y| \left(R_{i,j+\frac{1}{2}}^y \right)^\top [V_{i,j+\frac{1}{2}}], \end{aligned} \quad (4.54)$$

where $\tilde{F}_{i+\frac{1}{2},j}$, $\tilde{G}_{i,j+\frac{1}{2}}$ are any pair of consistent, energy preserving fluxes satisfying (4.48), V is the vector of energy variables, $R_{i+\frac{1}{2},j}^x$ and $R_{i,j+\frac{1}{2}}^y$ are defined in (4.52) and

$$\begin{aligned} |\Lambda_{i+\frac{1}{2},j}^x| &= \text{diag} \left\{ \left| \bar{u}_{i+\frac{1}{2},j} - \sqrt{g\bar{h}_{i+\frac{1}{2},j}} \right|, \left| \bar{u}_{i+\frac{1}{2},j} \right|, \left| \bar{u}_{i+\frac{1}{2},j} + \sqrt{g\bar{h}_{i+\frac{1}{2},j}} \right| \right\}, \\ |\Lambda_{i,j+\frac{1}{2}}^y| &= \text{diag} \left\{ \left| \bar{v}_{i,j+\frac{1}{2}} - \sqrt{g\bar{h}_{i,j+\frac{1}{2}}} \right|, \left| \bar{v}_{i,j+\frac{1}{2}} \right|, \left| \bar{v}_{i,j+\frac{1}{2}} + \sqrt{g\bar{h}_{i,j+\frac{1}{2}}} \right| \right\}. \end{aligned}$$

The above flux is clearly consistent. A simple generalization of Theorem 4.2 shows that the finite volume scheme based on fluxes (4.54) is energy stable. We skip the details of this estimate as it is a straightforward generalization of (4.41).

We denote the two-dimensional entropic Roe-type scheme with energy preserving fluxes (4.49), (4.54) as the ERoe scheme and compare it with standard Rusanov and Roe schemes. We consider the two-dimensional cylindrical dam-break problem (4.50) and compute with the standard Roe and ERoe schemes on a uniform 100×100 mesh. Figure 4.14 shows

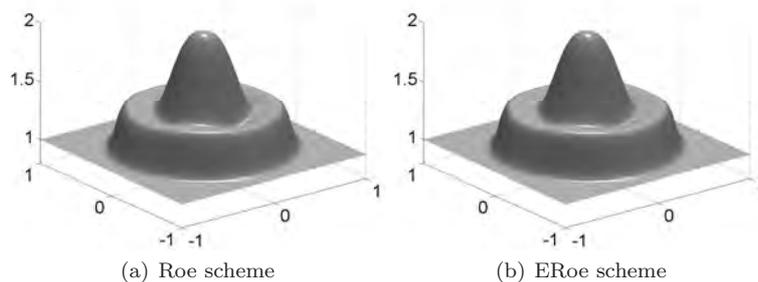


Fig. 4.14. Approximate heights for the cylindrical dam-break problem at time $t = 0.2$ computed on a uniform 100×100 mesh with the Roe and ERoe schemes.

that both the standard Roe and ERoe schemes approximate the solution rather well, but with smearing at the outward shock and the rarefaction. As with one-dimensional dam-break problems, the differences between the two schemes are very small. However, as in Section 4.2, we can find several examples where the ERoe scheme is stable, whereas the standard Roe scheme is unstable.

As in Section 4.2.7, we have computed the cost with each scheme, and we present the normalized run-times (unit time is set to 0.0475 seconds) with respect to relative errors in height in Table 4.4. As shown in this table, the Roe and ERoe schemes have approximately the same computational cost and are considerably more efficient than the Rusanov scheme. Given the fact that the ERoe scheme is as accurate, has the same computational cost as the Roe scheme but is provably energy-stable, it seems preferable to the Roe scheme.

Relative height error (in %)	4	2	1
Rusanov	1.59	32.1	404.21
Roe	1	20.99	229.47
ERoe	1.03	21.17	231.57

Table 4.4. *Normalized run-times for the Rusanov, Roe and ERoe schemes on the two-dimensional cylindrical dam-break problem with three different levels of relative error in height.*

Numerical experiment #9: A physical dam-break problem

So far, we have considered model problems in our numerical experiments. In order to demonstrate the robustness of our approach, we will consider a more challenging two-dimensional dam-break problem with a physically realistic set up. This problem was first studied in Fennema and Chaudhery (1990) and was also considered in Tadmor and Zhong (2008), Chertock and Kurganov (2004).

The geometry of the problem and the initial conditions are specified in Figure 4.15. As shown in this figure, we consider a basin of 1400×1400

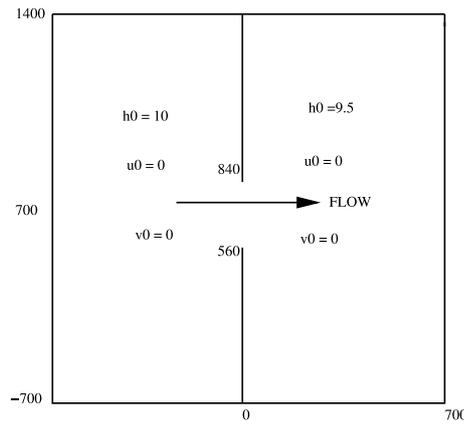


Fig. 4.15. Set-up and initial conditions of the physical dam-break problem

m^2 with a dam in the middle. The walls of the basin are solid and frictionless and the bottom is assumed to be flat. The walls are reflective with initial water level at 10 m and tail water level of 9.5 m . At $t = 0$, the central part of the dam fails and water is released downstream through a breach, as shown in Figure 4.15. We set acceleration due to gravity to 9.8ms^{-2} . The boundary treatment is similar to the one used in Tadmor

and Zhong (2008). We compute approximate solutions of this problem with three different schemes: the energy preserving EEC scheme, the EEC scheme together with eddy viscosity of $\nu = 10m^2s^{-1}$ and the energy stable ERoe scheme. We used a uniform 100×100 mesh with a second-order RK2 method for time integration at a CFL number of 0.45. The results are displayed in Figure 4.16. The results are along expected lines.

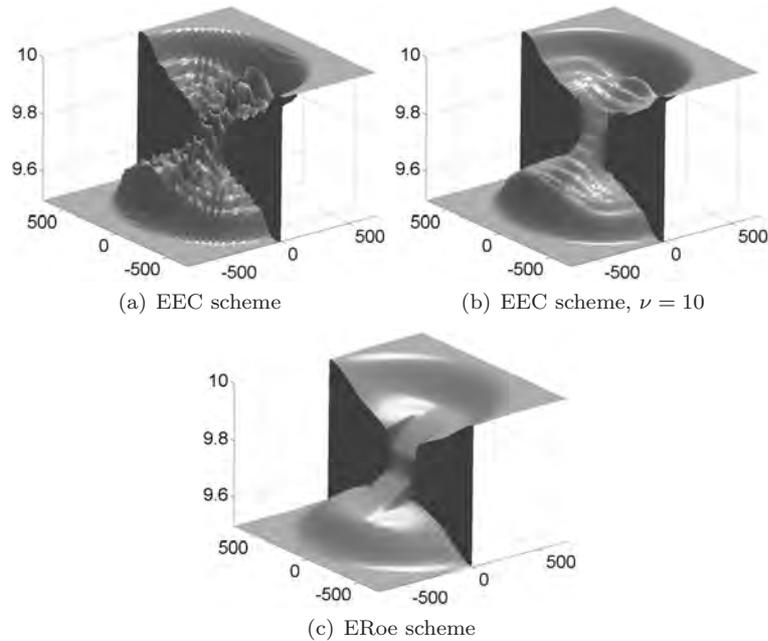


Fig. 4.16. Approximate heights for the physical dam-break problem at time $t = 50$ computed on a uniform 100×100 mesh with EEC scheme together with eddy viscosity and with the ERoe scheme.

The EEC schemes produce oscillations, but the basic flow features are computed quite accurately. In particular, the circular shock wave is resolved quite sharply.

The addition of eddy viscosity reduces the oscillations considerably. However, some oscillations are still present, indicating an under-resolved computation. When we combine the EEC fluxes (4.49) with the Roe-type entropy variables based numerical diffusion operator (4.54) to obtain the ERoe scheme, we observe that all the oscillations are removed. However, the shocks and other wave fronts are smeared. These results

are very similar to those obtained in Tadmor and Zhong (2008), Chertock and Kurganov (2004) and confirm the robustness of our approach.

4.3.4 Numerical Experiment #10: Advection of Vorticity

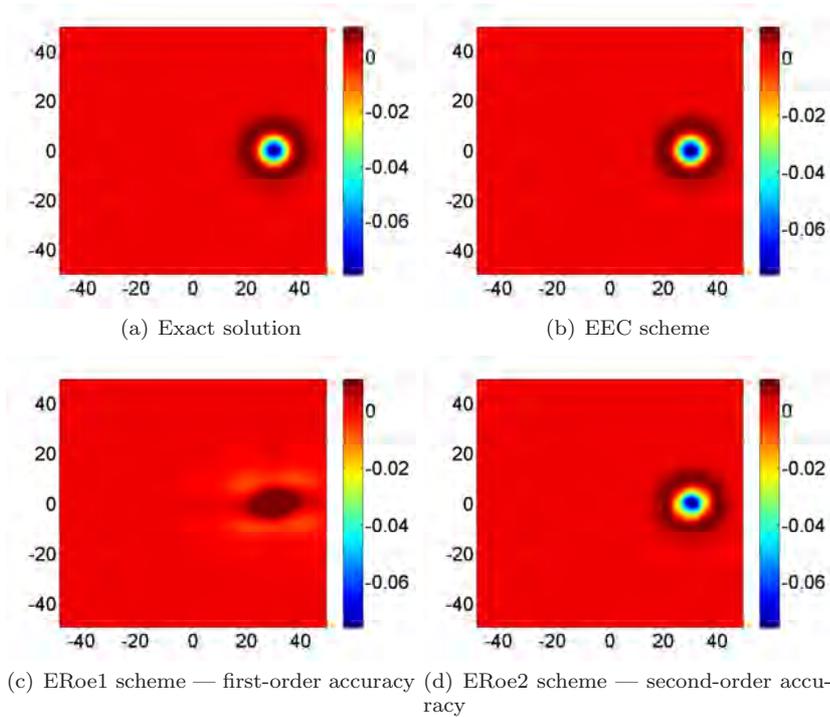


Fig. 4.17. Discrete vorticity at time $t = 100$ on a 200×200 mesh for numerical experiment 6: Vortex advection with EEC and ERoe schemes.

All the numerical experiments presented so far involved the formation and propagation of shocks. Energy preserving schemes led to oscillations near the shock on account of energy transfer to the small scales in the problem. We needed to introduce either eddy viscosity or numerical diffusion to remove these oscillations. Another interesting object of study, particularly for the two-dimensional form of the shallow water equations (4.2), is the vorticity. We define the vorticity as $\omega = v_x - u_y$, with u and v being the velocities. It is well known (Arakawa and Lamb (1977))

that the vorticity satisfies the following advection equation,

$$\omega_t + (u\omega)_x + (v\omega)_y = 0. \quad (4.55)$$

This identity is valid only for the smooth solutions of (4.2). Many papers, e.g., Arakawa and Lamb (1981), have dealt with the question of designing numerical schemes that satisfy a discrete version of the vorticity advection (4.55). The vorticity errors generated by a scheme discretizing (4.2) are considered a key component of its overall performance, particularly for meteorological applications. It was speculated in Tadmor and Zhong (2008) that energy preserving schemes for the shallow water equations might generate large vorticity errors. Hence, we test the energy preserving EEC scheme in a numerical experiment dealing with the advection of a vortex.

We adapt a standard test case (see Ismail and Roe (2005) and references therein) for the isentropic Euler equations to the shallow water equations. One can check that the following functions,

$$\begin{aligned} h(x, y, t) &= 1 - \frac{c_1^2}{4c_2g} e^{2f} \\ u(x, y, t) &= M + c_1(y - y_0)e^f \\ v(x, y, t) &= -c_1(x - x_0 - Mt)e^f, \end{aligned}$$

where

$$f = f(x, y, t) = -c_2 \left((x - x_0 - Mt)^2 + (y - y_0)^2 \right),$$

are a smooth solution to the shallow water equations (4.2) for any choice of constants M, c_1, c_2, x_0 and y_0 . For our numerical experiment, we consider the above solution at time $t = 0$ as our initial data and choose $M = 0.5$, $g = 1$, perturbation coefficients $(c_1, c_2) = (-0.04, 0.02)$ and starting position $(x_0, y_0) = (-20, 0)$. The exact solution is a vortex moving at a constant velocity in the x -direction. This test case has been considered in many papers in the literature and standard schemes have been found to generate unacceptably large vorticity errors. We compute on a domain $[-50, 50] \times [-50, 50]$ and show the approximate solutions in Figure 4.17.

We compute the standard discrete vorticity given by

$$\omega_{i,j} = \frac{v_{i+1,j} - v_{i-1,j}}{2\Delta x} - \frac{u_{i,j+1} - u_{i,j-1}}{2\Delta y},$$

and plot the discrete vorticity on a uniform 200×200 mesh with EEC and ERoe schemes at time $t = 100$. The time integration is performed

with an RK2 method at a CFL number of 0.45. The vorticity of the exact solution is plotted for comparison. As shown in the figure, the EEC scheme does a remarkable job of resolving the solution at this fairly coarse mesh with $\Delta x = \Delta y = 0.5$ and a long period of time with final time $t = 100$. There are no oscillations whatsoever, and the shape of the vortex is retained. This should be contrasted with the oscillations that a EEC scheme generates near shocks. Similarly, the performance of the EEC scheme is considerably better than the ERoe scheme. As seen in Figure 4.17(c), the structure of the vortex is destroyed in the first-order ERoe scheme. Such behavior is expected from standard schemes (Ismail and Roe (2005)). A possible explanation lies in the order of accuracy, as the EEC scheme is formally second-order accurate, whereas the ERoe1 scheme is restricted to first-order of accuracy. We extended the ERoe scheme to second-order accuracy by applying minmod limiters, as in Section 4.2.8; the result of ERoe2 is shown in Figure 4.17(d). The rendition is now much more accurate, but the increased accuracy comes at the cost of small spurious oscillations that make the vortex appear non-symmetric.

We believe that the order alone does not explain the good performance of the EEC scheme. One reason for its performance may lie in the preservation of energy, illustrated in Figure 4.18. The figure shows that the energy error is very low, and this may be one of the reasons that the scheme preserves the structure of the vortex.

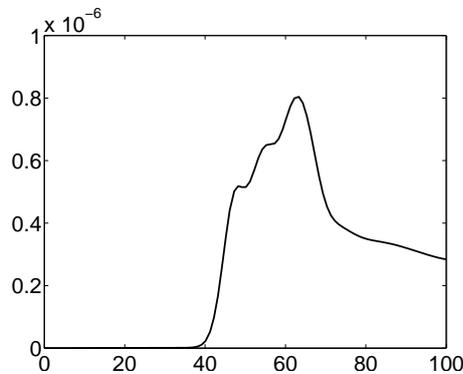


Fig. 4.18. Energy vs. time for EEC scheme in numerical experiment 6 with 200×200 mesh and RK2 method with $CFL = 0.45$.

This experiment illustrates that the EEC scheme is very robust for smooth solutions and needs no eddy viscosity or numerical diffusion

for stability. Furthermore, it advects vorticity without destroying the vortical structures in the problem.

4.4 Conclusion

We considered the shallow water equations in one and two space dimensions. Our aim was to design finite volume schemes for these equations that either preserve energy or dissipate it in the right way. We considered three different energy preserving schemes in one dimension – the PEC scheme proposed in Tadmor (2003), Tadmor and Zhong (2008), an explicit evaluation of the scheme proposed in Tadmor (1987), and a new explicit energy preserving scheme, inspired by the entropy preserving scheme for the Euler equations in Roe (2006). The new scheme, called the EEC scheme, is very easy to implement and is robust. We compared the three schemes in numerical experiments and observed that the PEC and EEC schemes behave in a similar manner. The main advantage of the EEC scheme over the PEC scheme is its low computational cost: about three times faster in one dimension and about six times faster in two dimensions. The low cost coupled with the ease of implementation makes the EEC scheme an ideal energy preserving scheme for the shallow water equations.

The energy preserving schemes lead to oscillations at the mesh scale due to the absence of diffusive mechanisms. Addition of eddy viscosity by using suitable discretization of the viscous terms leads to energy stable discretizations of the shallow water equations. Eddy viscosity dramatically reduces the oscillations, particularly on resolved meshes. However, oscillations might remain on under-resolved meshes.

Computations on under-resolved meshes or for the inviscid version of the equations require adding suitable artificial diffusion operators. We proposed a novel numerical flux of the Roe type that is based on using an energy preserving symmetric flux together with a numerical diffusion based on energy variables. The diffusion operators need to be scaled suitably, and the resulting scheme is energy-stable without any additional fixes. The resulting scheme, termed an ERoe scheme, is very robust. It has the same accuracy and a similar computational cost as the standard Roe scheme, but is more stable.

All the above points are demonstrated in a series of numerical experiments in both one and two spatial dimensions. We believe that this approach of using energy (entropy) preserving schemes together with suitable physical or numerical diffusions operators is a practical,

cost-effective and stable approach for computing flows. We will aim to extend this approach to more complicated models like the Euler equations of gas dynamics, incompressible flows and equations of magneto-hydrodynamics in forthcoming papers.

Acknowledgment. The work on this paper was started when S. M. visited the Center of Scientific Computation and Mathematical Modelling (CSCAMM) and he thanks CSCAMM and all its members for the excellent hospitality and facilities. The research of E. T. was supported in part by NSF grant 07-07949 and ONR grant N00014-91-J-1076.

References

- A. Arakawa (1966), ‘Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow’, *J. Comput. Phys.* **1**, 119–143.
- A. Arakawa and V. Lamb (1977), ‘Computational design of the basic dynamical process of the UCLA general circulation model’, *Meth. Comput. Phys.* **17**, 173–265.
- A. Arakawa and V. Lamb (1981), ‘A potential enstrophy and energy conserving scheme for the shallow water equations’, *Mont. Weat. Rev.* **109**, 18–36.
- A. Chertock and A. Kurganov (2004), ‘On a hybrid finite-volume-particle method’, *M2AN Math. Model. Num. Anal.* **38**, 1971–1991.
- B. Cockburn, C. Johnson, C.-W. Shu and E. Tadmor (1998), *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations*, C.I.M.E. course held in Cetraro, Italy, June 1997 (A. Quarteroni ed.), Lecture notes in Mathematics **1697**, Springer Verlag.
- C. Dafermos (2000), *Hyperbolic Conservation Laws in Continuum Physics*, Springer, Berlin.
- R. Fennema and M. Chaudhry (1990), ‘Explicit methods for the 2d-transient free surface flows’ *J. Hydraul. Eng. ASCE.* **116**, 1013–1034.
- S. Gottlieb, C.-W. Shu and E. Tadmor (2001), ‘High order time discretizations with strong stability properties’, *SIAM. Review* **43**, 89–112.
- J. Goodman and P. Lax (1988), ‘On dispersive-diffusive schemes I’, *Comm. Pure. Appl. Math.* **41**, 591–613.
- T. Hou and P. Lax (1991), ‘Dispersive approximations in fluid dynamics’, *Comm. Pure. Appl. Math.* **44**, 1–40.
- T. Hughes, L. Franca and M. Mallet (1986), ‘A new finite element formulation for CFD I: Symmetric forms of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics’, *Comp. Meth. Appl. Mech. Eng.* **54**, 223–234.
- F. Ismail and P. Roe (2005), ‘Towards a vorticity preserving second order finite volume scheme solving the Euler equations’, *Proc. of AIAA conference on CFD*, Toronto, Canada.
- K. Khalfallah and A. Lerat (1989), ‘Correction d’entropie pour des schémas numériques approchant un système hyperbolique’, *C. R. Acad. Sci. Paris Sér. II* **308**, 815–820.
- A. Kurganov and E. Tadmor (2002), ‘New high resolution central schemes for non-linear conservation laws and convection-diffusion equations’, *J. Comput. Phys* **160**, 241–282.

- P. LeFloch, J. Mercier and C. Rohde (2002), ‘Fully discrete entropy conservative schemes of arbitrary order’, *SIAM J. Numer. Anal.* **40**, 1968–1992.
- R. LeVeque (2002), *Finite Volume Methods for Hyperbolic Problems*, Cambridge University Press, Cambridge.
- P. Roe (2006), ‘Entropy conservative schemes for Euler equations’, talk at *HYP 2006, Lyon, France*, unpublished. Lecture available from <http://math.univ-lyon1.fr/~hyp2006>.
- E. Tadmor (1984), ‘Numerical viscosity and entropy conditions for conservative difference schemes’, *Math. Comp.* **43**, 369–381.
- E. Tadmor (1986), ‘Entropy conservative finite element schemes’ in *Numerical Methods for Compressible Flows - Finite Difference Element and Volume Techniques*, Proc. annual meeting of the AMSE (T. E. Tezduyar and T. J. R. Hughes, eds.), v. **78**, 149–158.
- E. Tadmor (1987), ‘The numerical viscosity of entropy stable schemes for systems of conservation laws, I.’, *Math. Comp.* **49**, 91–103.
- E. Tadmor (2002), ‘From semi-discrete to fully discrete: stability of Runge-Kutta schemes by the energy method. II.’, in *Collected Lectures on the Preservation of Stability under Discretization*, Lecture Notes Colorado State Univ. Conference, Fort Collins, CO, 2001 (D. Estep and S. Tavener, eds.), Proc. Appl. Math. **109**, SIAM, 25–49.
- E. Tadmor (2003), ‘Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems’, *Acta Numerica* **12**, 451–512.
- E. Tadmor and W. Zhong (2006), ‘Entropy stable approximations of Navier-Stokes equations with no artificial numerical viscosity’, *J. Hyperbolic. Differ. Equ.* **3**, 529–559.
- E. Tadmor and W. Zhong (2008), ‘Energy preserving and stable approximations for the two-dimensional shallow water equations’, in *Mathematics and Computation: A Contemporary View*, Proc. of the third Abel symposium, Ålesund, Norway, Springer, 67–94.