Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

P.J. van der Houwen, B.P. Sommeijer

Reduction of dispersion in hyperbolic difference schemes by adapting the space discretization

Department of Numerical Mathematics Report NM-R8519 September

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Reduction of Dispersion in Hyperbolic Difference Schemes by Adapting the Space Discretization

P.J. van der Houwen & B.P. Sommeijer

A fourth-order accurate difference scheme for systems of hyperbolic equations is presented. The dispersion in this scheme can be reduced if it is known in advance in which region the frequencies of the dominant Fourier components are located. All method parameters are explicitly expressed in terms of the bounds on the dominating frequencies. The performance of the method is illustrated by an application to the shallow water equations.

1980 Mathematics Subject Classification: Primary: 65M20; Secondary: 76B15
Key Words & Phrases: hyperbolic equations, difference schemes, spatial discretization, dispersion, shallow water equations.
Note: This report is will be submitted for publication elsewhere.

1. Introduction

In [3] we proposed a fourth-order difference method for solving hyperbolic systems of the form

$$\frac{\partial \mathbf{w}}{\partial t} = A(\mathbf{w})\frac{\partial}{\partial x}\mathbf{w} + B(\mathbf{w})\frac{\partial}{\partial y}\mathbf{w}, \tag{1.1}$$

where A and B are symmetric matrices. This method is particularly suited to solve problems where: (i) A and B are slowly varying with $\mathbf{w}$; (ii) it is known in advance that the exact solution mainly consists of one-dimensional waves of the form $a\,exp[i(\alpha t+\omega_x(x+cy)]$ with c constant and (α,ω_x) lying in a given region $[\underline{\alpha},\overline{\alpha}]\chi[\underline{\omega}_x,\overline{\omega}_x]$. This was achieved by constructing a special Runge-Kutta time integrator with small phase errors (dispersion) with respect to these one-dimensional waves.

For *ordinary* differential equations, a considerable amount of work is done in constructing methods with small phase errors when integrating exponentials of the form $a\,exp(i\alpha t)$. We mention a few recent papers: [1],[2],[4],[5] and [6]. Sofar, the extension of such an approach to *partial* differential equations with periodic solutions has received little attention.

In this paper we further develop the method proposed in [3] by including special spatial discretization formulas for $\partial/\partial x$ and $\partial/\partial y$. In Section 2, the basic results of [3] are summarized and the dispersion of the difference method is introduced. Section 3 presents two strategies for reducing the dispersion in a prescribed region $\mathcal{R} = [\underline{\alpha}\Delta t,\overline{\alpha}\Delta t]\chi[\underline{\omega}_x\Delta x,\overline{\omega}_x\Delta x]$. In Sections 4 and 5 these strategies are applied to an 8-point and a 16-point space-discretization formula, respectively. Finally, Section 6 presents numerical results for the shallow water equations.

2. Preliminaries

Using the method of lines approach we first discretize the spatial derivative operators $\partial/\partial x$ and $\partial/\partial y$ by a symmetric difference operator:

$$\frac{\partial}{\partial x} \approx D_x := \frac{1}{2\Delta x}\sum_{l=0}^{k}\sum_{j=1}^{k} \zeta_j^{(l)}[E_x^j - E_x^{-j}][E_y^l + E_y^{-l}], \tag{2.1}$$

E_x and E_y being shift operators in the x and y direction, respectively, and a similar approximation for $\partial/\partial y$; next, we integrate the resulting system of ordinary differential equations (ODEs)

Report NM-R8519
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

$$\frac{d\mathbf{W}}{dt} = [A(\mathbf{W})D_x + B(\mathbf{W})D_y]\mathbf{W} \tag{2.2}$$

by a four-point Runge-Kutta method generated by the Butcher array

$$\begin{array}{c|cccc} 0 & 0 & & & \\ 1/2 & 1/2 & & & \\ 1/2 & 0 & 1/2 & & \\ 1 & 0 & 0 & 1 & \\ \hline & b_1 & b_2 & 1-b_2-2b_1 & b_1 \end{array} \tag{2.3}$$

where b_1 and b_2 are free parameters. The following theorem was proved in [3].

THEOREM 2.1. *Let $\partial/\partial x$ and $\partial/\partial y$ be discretized by (2.1), both using the same weights $\zeta_j^{(l)}$, and let (2.2) be discretized by (2.3). Furthermore, let Δx and Δy be both $O(\Delta t)$ as $\Delta t \to 0$. Then the difference scheme is fourth-order accurate if*

$$\sum_{l,j}^{k} j\zeta_j^{(l)} = \frac{1}{2}, \quad \sum_{l,j}^{k} j^3\zeta_j^{(l)} = O(\Delta^2 t), \quad \sum_{l,j}^{k} jl^2\zeta_j^{(l)} = O(\Delta^2 t), \tag{2.4a}$$

$$b_1 = \frac{1}{6} + O(\Delta^2 t), \quad b_2 = \frac{1}{3} + O(\Delta^2 t). \quad \square \tag{2.4b}$$

The conditions (2.4a) and (2.4b) respectively correspond to the consistency of the space discretization and the time discretization.

We will exploit the freedom left by the order conditions (2.4) for improving the accuracy of the numerical solution with respect to certain dominant solution components. Our approach is based on the following theorem, the proof of which can be found in [3].

THEOREM 2.2. *Let A and B be constant matrices and let the exact solution contain a dominant component $\mathbf{w}_0 = \mathbf{a}\exp[i(\alpha t + \omega_x x + \omega_y y)]$, where $\mathbf{a}$ is an eigenvector of the matrix $\omega_x A + \omega_y B$ with eigenvalue α. If $\Delta y = |\omega_x|\Delta x/|\omega_y|$ then, for $t = n\Delta t$, the numerical error corresponding to the component $\mathbf{w}_0$ is given by*

$$[e^{in\nu} - R_4^n(i\nu\delta(\mu))]\mathbf{w}_0(0, x_j, y_l), \quad (x_j, y_l) \in \Omega_\Delta, \tag{2.5}$$

where

$$\mu := \omega_x \Delta x, \quad \nu := \alpha \Delta t,$$

$$R_4(z) := 1 + z + \frac{1}{2}z^2 + \frac{1}{4}(1-b_2)z^3 + \frac{1}{4}b_1 z^4,$$

$$\delta(\mu) := 2\sum_{j,l}^{k} \xi_j^{(l)} \frac{\sin(j\mu)\cos(l\mu)}{\mu}. \quad \square$$

In this paper we concentrate on the reduction of the numerical error corresponding to particular Fourier components in the initial function characterized by the frequencies $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, $\omega_x \in [\underline{\omega}_x, \overline{\omega}_x]$ and $\omega_y \in [\underline{\omega}_y, \overline{\omega}_y]$. It will be assumed that Δy is chosen according to

$$\Delta y = \left|\frac{\underline{\omega}_x + \overline{\omega}_x}{\underline{\omega}_y + \overline{\omega}_y}\right| \Delta x. \tag{2.6}$$

If Δx and Δy differ much in magnitude one may consider a rotation of the coordinate axes in order to balance Δx and Δy.

Our first step will be *exponential fitting*, that is we determine b_1 and b_2 in (2.3) such that

$$R_4(i\nu_0\delta(\mu_0)) = e^{i\nu_0}, \tag{2.7}$$

where the fitting point (μ_0,ν_0) is chosen in $\mathcal{R} := \{(\mu,\nu) | \underline{\mu} \leq \mu \leq \bar{\mu}, \underline{\nu} \leq \nu \leq \bar{\nu}\}$ with obvious definition of $\underline{\mu}$, $\bar{\mu}$, $\underline{\nu}$ and $\bar{\nu}$. Together with the order conditions (2.4a) exponential fitting implies fourth-order accuracy in time [3].

THEOREM 2.3. *Let the discretization function $\delta(\mu)$ satisfy condition (2.7) and let (2.4a) be satisfied. Then condition (2.4b) is fulfilled.* □

The fitting condition (2.7) determines the Runge-Kutta parameters b_1 and b_2 as functions of the fitting point (μ_0,ν_0).

Our next step is the reduction of the numerical error components corresponding to points $(\mu,\nu) \in \mathcal{R}$ by a judicious choice of the parameters μ_0,ν_0 and $\xi_j^{(l)}$ under the constraints (2.4a) and (2.7). In particular, we will reduce the magnitude of the *dispersion* of the scheme, defined by

$$\phi(\mu,\nu) := \nu - arg(R_4(i\nu\delta(\mu))), \tag{2.8}$$

for $(\nu,\mu) \in \mathcal{R}$ (notice that $\phi(\mu_0,\nu_0)=0$).

In the considerations above, the matrices A and B are assumed to be constant. Formally, however, these considerations can be extended to nonlinear equations where A and B depend on $\mathbf{W}$; then, Theorem 2.2 should be interpreted locally, and consequently, the Runge-Kutta parameters (b_1,b_2), the fitting point (μ_0,ν_0), the region of dominant frequencies $\mathcal{R}$ and the discretization weights $\xi_j^{(l)}$ depend on the solution so that they are to be calculated during the integration process.

To conclude this summary, we remark that the stability behaviour of these exponentially fitted methods is almost identical to that of the classical fourth-order method; in particular, it can be shown [3] that the imaginary stability boundary is given by $2\sqrt{2} - O(\nu_0^2\delta^2(\mu_0))$.

3. REDUCTION OF DISPERSION

It is our purpose to reduce the value of

$$\|\phi\|_{\mathcal{R}} := \max_{\mathcal{R}} |\phi(\nu,\mu)| \tag{3.1}$$

by a suitable choice of μ_0, ν_0 and $\xi_j^{(l)}$. Instead of minimizing $\|\phi\|_{\mathcal{R}}$ by a numerical search, we will look for less expensive methods to reduce its value by analytical means.

Since $\phi(\mu_0,\nu_0) = 0$, we want to keep ϕ small if μ and ν move away from μ_0 and ν_0, respectively. In order to reduce the effect of perturbations of μ_0 we will require that $\delta(\mu)$ has a small range on the interval $[\underline{\mu},\bar{\mu}]$; the effect of perturbing ν_0 will be reduced by requiring $\partial\phi/\partial\nu = 0$ at (μ_0,ν_0). We will first consider this latter type of perturbation.

3.1. Reduction of dispersion due to perturbations of the time frequency

We write $\phi(\mu,\nu)$ in the form

$$\phi(\mu,\nu) = \nu - a(\nu\delta(\mu)),\ a(z) := \arctan\left[\frac{z(1-(1-b_2)z^2/4)}{1-\frac{1}{2}z^2+b_1z^4/4}\right]. \tag{3.2}$$

Hence,

$$\frac{\partial\phi}{\partial\nu} = 1 - a'(\nu\delta(\mu))\,\delta(\mu).$$

By virtue of the fitting condition (2.7) it can be shown that

$$a'(\nu_0\delta(\mu_0)) = \frac{4\sin\nu_0 - \sin\nu_0\cos\nu_0 - 2\nu_0\delta(\mu_0)\cos\nu_0 - \nu_0^2\delta^2(\mu_0)\sin\nu_0}{\nu_0\delta(\mu_0)}.$$

Thus, the condition $\phi_\nu(\mu_0,\nu_0) = 0$ leads us to the equation

$$\nu_0^2\sin\nu_0\delta^2(\mu_0) + 2\nu_0\cos\nu_0\delta(\mu_0) + \nu_0 - 4\sin\nu_0 + \sin\nu_0\cos\nu_0 = 0.$$

For given ν_0 we find that $\delta(\mu_0)$ is given by

$$\delta(\mu_0) = \delta_0 := \frac{\cos\nu_0}{\nu_0\sin\nu_0}\left[\sqrt{1+\frac{\sin\nu_0(4\sin\nu_0 - \nu_0 - \sin\nu_0\cos\nu_0)}{\cos^2\nu_0}} - 1\right] \tag{3.3}$$

$$\approx 1 - \frac{1}{120}\nu_0^4.$$

Of course, this is not necessarily possible for an arbitrarily prescribed discretization function $\delta(\mu)$. We illustrate this by means of the conventional fourth-order 8-point discretization where all $\xi_j^{(l)}$ in (2.1) vanish except for

$$\xi_1^{(0)} = \frac{2}{3},\quad \xi_2^{(0)} = -\frac{1}{12}. \tag{3.4}$$

The corresponding discretization function is given by

$$\delta(\mu) = \frac{\sin\mu}{3\mu}(4-\cos\mu) \approx 1 - \frac{1}{30}\mu^4.$$

A comparison with (3.3) reveals that this function can only satisfy (3.3) if $\mu_0 \approx \nu_0 / \sqrt{2}$. Since also $(\mu_0,\nu_0) \in \mathcal{R}$, this is not necessarily true. Thus, in constructing a suitable discretization function we add to the conditions (2.4a), the condition (3.3). Notice that (3.3) is compatible with the order equations (2.4a).

3.2. Reduction of dispersion due to perturbations of the spatial frequency

First of all we observe that the parameters $\xi_j^{(l)}$ are not independently free parameters. To see this we write $\delta(\mu)$ in the form

$$\delta(\mu) = 2\frac{\sin\mu}{\mu}\sum_{l=0}^{k}\sum_{j=1}^{k}\xi_j^{(l)}U_{j-1}(\cos\mu)T_l(\cos\mu), \tag{3.5}$$

where T_l and U_{j-1} represent Chebyshev polynomials of the first and second kind. Defining the polynomial

$$P_r(z) := \sum_{l=0}^{k}\sum_{j=1}^{k}\xi_j^{(l)}U_{j-1}(z)T_l(z) = \sum_{i=0}^{r}p_i z^i \tag{3.6}$$

we can represent $\delta(\mu)$ in the form

$$\delta(\mu) = 2\frac{\sin\mu}{\mu}P_r(\cos\mu). \tag{3.7}$$

The coefficients p_i are linear expressions in terms of the $\xi_j^{(l)}$. The value of r is given by

$$r = \max_{\xi_j^{(l)}\neq 0}(l+j-1). \tag{3.8}$$

Obviously, $r \leqslant 2k-1$, so that we have at most $2k$ independent, free parameters whereas there are

k^2+k parameters $\xi_j^{(l)}$.

We want to minimize the range of $\delta(\mu)$ on $[\underline{\mu},\overline{\mu}]$ under the constraints (2.4a) and (3.3). The first condition in (2.4a) is equivalent to

$$\delta(0) = 1. \tag{3.9}$$

We will impose (3.3) and (3.9) on $\delta(\mu)$ and we will check the remaining conditions of (2.4a) later on (we observe that (3.3) implies the first condition of (2.4a)).

It seems reasonable to choose $\delta(\mu)$ such that $\delta(\mu)-\delta_0$ has as many zeros as possible in $[\underline{\mu},\overline{\mu}]$. Since there are $r+1$ free coefficients p_i, we expect, in view of (3.9), that we can assign r distinct zeros $\mu_1,\mu_2,\ldots,\mu_r$ to $\delta(\mu)-\delta_0$. Observing that $\delta(\mu)$ is an even function and using Lagrange's interpolation formula, we may represent $\delta(\mu)$ in the form

$$\delta(\mu) = \delta_0 + \pi_r(\mu^2)\left[\frac{1-\delta_0}{\pi_r(0)} + \frac{\delta^{(r)}(\theta(\mu^2))-\delta^{(r)}(\theta(0))}{r!}\right], \tag{3.5'}$$

where

$$\pi_r(\mu^2) := (\mu^2-\mu_1^2)(\mu^2-\mu_2^2)\cdots(\mu^2-\mu_r^2)$$

and where $\theta(\mu^2)$ assumes values in the interval $[\min(\underline{\mu}^2,\mu^2), \max(\overline{\mu}^2,\mu^2)]$. This representation suggests choosing for $\pi_r(\mu^2)$ a minimax polynomial on the interval $[\underline{\mu}^2,\overline{\mu}^2]$. This leads us to choose the abscissas $\{\mu_j^2\}$ equal to the zeros $\{z_j\}$ of the shifted Chebyshev polynomial

$$T_r\left[\frac{2\mu^2-\overline{\mu}^2-\underline{\mu}^2}{\overline{\mu}^2-\underline{\mu}^2}\right]. \tag{3.10}$$

The resulting (linear) system for the coefficients p_i is given by

$$\begin{cases} \sum_{i=0}^{r} p_i = \frac{1}{2}, \\ 2\,\frac{\sin\sqrt{z_j}}{\sqrt{z_j}} \sum_{i=0}^{r} p_i \cos^i\sqrt{z_j} = \delta_0, \quad j = 1,2,\ldots,r. \end{cases} \tag{3.11}$$

The zeros z_j are given by

$$z_j := \tfrac{1}{2}(\overline{\mu}^2+\underline{\mu}^2)+\tfrac{1}{2}(\overline{\mu}^2-\underline{\mu}^2)\cos\left(\frac{2j-1}{2r}\pi\right), \quad j = 1,2,\ldots,r. \tag{3.12}$$

Since all z_j in $\{z_j\}$ are distinct, the system (3.11) can be solved providing us with the more or less optimal coefficients p_i and, by virtue of (3.6), with the parameters $\xi_j^{(l)}$. However, these parameters are still functions of $\nu_0 \in [\underline{\nu},\overline{\nu}]$, because of the constraint (3.3), and of $\mu_0 \in \{\sqrt{z_j}\}$. By a numerical search we may determine the optimal fitting point (μ_0,ν_0).

We remark that in cases where $\underline{\nu} \approx \overline{\nu}$ we may drop the constraint (3.3) giving us an additional parameter for minimizing the range of $\delta(\mu)$. Proceeding as above we obtain instead of (3.11) the system

$$\begin{cases} \sum_{i=0}^{r} p_i = \frac{1}{2}, \\ \frac{\sin\sqrt{z_j}}{\sqrt{z_j}} \sum_{i=0}^{r} p_i \cos^i\sqrt{z_j} = \frac{\sin\sqrt{z_1}}{\sqrt{z_1}} \sum_{i=0}^{r} p_i \cos^i\sqrt{z_1}, \; j = 2,3,\ldots,r+1, \end{cases} \tag{3.13}$$

where the z_j are defined according to (3.12), with r replaced by $r+1$. Here, $\nu_0 \in [\underline{\nu},\overline{\nu}]$ and $\mu_0 \in [\underline{\mu},\overline{\mu}]$. The optimal (μ_0,ν_0) may again be determined by a numerical search. In an actual computation, one

may decide to set $\nu_0 = (\bar{\nu}+\underline{\nu})/2$ in view of our assumption that $\underline{\nu} \approx \bar{\nu}$.

4. An 8-point Space-discretization Formula

Consider the difference operator

$$D_x = \frac{1}{\Delta x}[\xi_1^{(0)}(E_x - E_x^{-1}) + \xi_2^{(0)}(E_x^2 - E_x^{-2})] \tag{4.1}$$

and the operator D_y defined in a similar way using the same weights. Together they use 8 points, similar to the conventional fourth-order formula.

The corresponding discretization function $\delta(\mu)$ is given by (cf.(3.7))

$$\delta(\mu) = 2\frac{\sin\mu}{\mu}(p_0 + p_1\cos\mu), \tag{4.2}$$

where, according to (3.6)

$$p_0 = \xi_1^{(0)}, \quad p_1 = 2\xi_2^{(0)}. \tag{4.3}$$

Let us first determine p_0 and p_1 by solving (3.11) with $r=1$. Substituting the solution into (4.3) we find for $\xi_1^{(0)}$ and $\xi_2^{(0)}$ the expressions

$$\xi_1^{(0)} = \frac{\delta_0\mu_0 - \sin\mu_0\cos\mu_0}{2\sin\mu_0(1-\cos\mu_0)}, \quad \xi_2^{(0)} = \frac{1}{4} - \frac{1}{2}\xi_1^{(0)}, \tag{4.4}$$

where $\mu_0 = \sqrt{(\underline{\mu}^2 + \bar{\mu}^2)/2}$ and δ_0 is defined by (3.3).

It is easily verified that

$$\xi_1^{(0)} = \frac{2}{3} + \frac{\delta_0 - 1}{\mu_0^2} + O(\mu_0^2), \quad \xi_2^{(0)} = -\frac{1}{12} - \frac{\delta_0 - 1}{2\mu_0^2} + O(\mu_0^2),$$

which satisfy the order conditions (2.4a) provided that $\Delta x = O(\Delta t)$ as $\Delta t \to 0$. (we note that the order conditions (2.4b) are also satisfied by virtue of Theorem 2.3).

Next we solve the system (3.13) with $r=1$, to obtain

$$\xi_1^{(0)} = -\frac{1}{4}\frac{\sqrt{z_1}\sin 2\sqrt{z_2} - \sqrt{z_2}\sin 2\sqrt{z_1}}{\sqrt{z_1}\sin\sqrt{z_2}(1-\cos\sqrt{z_2}) - \sqrt{z_2}\sin\sqrt{z_1}(1-\cos\sqrt{z_1})}, \tag{4.5}$$

$$\xi_2^{(0)} = \frac{1}{4} - \frac{1}{2}\xi_1^{(0)},$$

where z_1 and z_2 are given by (3.12).

Since

$$\xi_1^{(0)} \approx \frac{2}{3}[1 - \frac{1}{5}(z_1+z_2)][1 + \frac{1}{4}(z_1+z_2)],$$

it follows that (2.4a) is satisfied provided $\Delta x = O(\Delta t)$ as $\Delta t \to 0$.

In Table 4.1 we have listed the gain factors $\|\phi_c\|_\infty / \|\phi_e\|_\infty$;here, ϕ_c corresponds to the standard fourth-order Runge-Kutta method ($b_1 = 1/6, b_2 = 1/3$) with the conventional discretization operators D_x and D_y ($\xi_1^{(0)} = 2/3$, $\xi_2^{(0)} = -1/12$), and ϕ_e corresponds to the exponentially fitted Runge-Kutta method (defined by (2.7)) with discretization weights (4.4). The optimal value of ν_0 was obtained by numerical minimization.

TABLE 4.1. Gain factors $\|\phi_c\|_\infty / \|\phi_e\|_\infty$ based on the weights (4.4)

(a) $\hat{\mu} = .25$, $\Delta\mu = 0$

$\hat{\nu} \setminus \Delta\nu$	5%	10%	20%	50%
.1	16603	4382	1229	301
.2	1200	323	95	29
.3	363	103	34	14
.4	224	67	24	12
.5	193	59	22	12

(b) $\hat{\nu} = .25$, $\Delta\nu = 0$

$\hat{\mu} \setminus \Delta\mu$	5%	10%	20%	50%
.1	10.4	5.3	2.8	1.5
.2	10.8	5.8	3.3	2.1
.3	10.9	5.9	3.5	2.2
.4	10.8	5.9	3.5	2.2
.5	10.7	5.8	3.5	2.2

The gain factors are given for a number of locations of the center $\hat{\mu}$ and $\hat{\nu}$ of the intervals $[\underline{\mu},\overline{\mu}]$ and $[\underline{\nu},\overline{\nu}]$, respectively. Furthermore, we have considered several uncertainty percentages $\Delta\mu$ and $\Delta\nu$ defined by

$$\Delta\mu := 100\,\frac{\overline{\mu}-\hat{\mu}}{\hat{\mu}}, \quad \Delta\nu := 100\,\frac{\overline{\nu}-\hat{\nu}}{\hat{\nu}}.$$

Table 4.2 represents the analogue of Table 4.1 when using (4.5) for defining the weights $\xi_j^{(0)}$. The optimal values of (μ_0,ν_0) were obtained numerically.

In these Tables the different features of both strategies are clearly demonstrated. In case of small μ-intervals (cf. the (a)-parts of both Tables), the first strategy (in which we required $\partial\phi/\partial\nu = 0$ at (μ_0,ν_0)) is superior. The second strategy, where we minimized the δ-range should be effective in cases where the μ-intervals are relatively large whereas the ν-intervals are small (cf. the (b)-parts of the Tables 4.1 and 4.2).

5. A 16-POINT SPACE-DISCRETIZATION FORMULA

Consider the discretization operator D_x with $k=2$ and $\xi_2^{(2)} = 0$. The corresponding discretization function is given by (3.7) with $r=2$:

$$\delta(\mu) = 2\,\frac{\sin\mu}{\mu}\,(p_0+p_1\cos\mu+p_2\cos^2\mu), \tag{5.1}$$

where

$$p_0 = \xi_1^{(0)}-\xi_1^{(2)}, \quad p_1 = 2\xi_2^{(0)}+\xi_1^{(1)}, \quad p_2 = 2(\xi_2^{(1)}+\xi_1^{(2)}). \tag{5.2}$$

TABLE 4.2. Gain factors $\|\phi_c\|_\infty / \|\phi_e\|_\infty$ based on the weights (4.5)

(a) $\hat{\mu} = .25$, $\Delta\mu = 0$

$\hat{\nu} \backslash \Delta\nu$	5%	10%	20%	50%
.1	10.3	5.6	3.4	3.0
.2	10.4	5.7	3.7	3.8
.3	10.9	6.4	4.7	5.5
.4	11.5	7.1	5.7	6.4
.5	12.0	7.6	6.4	6.7

(b) $\hat{\nu} = .25$, $\Delta\nu = 0$

$\hat{\mu} \backslash \Delta\mu$	5%	10%	20%	50%
.1	2144	549	145	29
.2	358	102	33	11
.3	260	77	27	9.9
.4	240	72	25	9.4
.5	230	69	24	8.9

Let us write

$$p_0 = \frac{2}{3} + q_0\Delta^2 t, \quad p_1 = -\frac{1}{6} + q_1\Delta^2 t, \quad p_2 = q_2\Delta^2 t. \tag{5.3}$$

If, on substitution into (3.11) or (3.13), we can find bounded values for q_i as $\Delta t \to 0$, then the equations (2.4a) are solved by

$$\xi_1^{(0)} = \frac{2}{3} + q_0\Delta^2 t, \quad \xi_2^{(0)} = -\frac{1}{12} + \frac{1}{2}q_1\Delta^2 t, \quad \xi_2^{(1)} = \frac{1}{2}q_2\Delta^2 t, \tag{5.4}$$
$$\xi_1^{(2)} = \xi_1^{(1)} = 0.$$

Let us first consider the system (3.11). Substitution of (5.3) yields

$$\begin{cases} \sum_{i=0}^{2} q_i = 0, \\ \sum_{i=0}^{2} q_i \cos^i \sqrt{z_j} = (\Delta t)^{-2}\left[\frac{1}{2}\delta_0 \dfrac{\sqrt{z_j}}{\sin\sqrt{z_j}} - \frac{2}{3} + \frac{1}{6}\cos\sqrt{z_j}\right] \\ \qquad\qquad = (\Delta t)^{-2}\left[\frac{1}{2}(\delta_0 - 1)(1 + \frac{1}{6}z_j) + O(z_j^2)\right], \ j = 1,2. \end{cases} \tag{5.5}$$

Since $\delta_0 = 1 + O(\Delta^2 t)$ and $z_j = O(\Delta^2 t)$, this system allows a bounded solution (q_0, q_1, q_2). Thus, (5.4) and (5.5) define a fourth-order accurate, 16-point discretization of $\partial/\partial x$ and $\partial/\partial y$, satisfying the system (3.11), for all $\nu_0 \in [\underline{\nu}, \overline{\nu}]$ and $\mu_0 \in \{\sqrt{z_j}\}$.

In the case of system (3.13) we obtain, on substitution of (5.3):

$$\begin{cases} \sum_{i=0}^{2} q_i = 0, \\ -\sum_{i=0}^{2} q_i \left[\dfrac{\sin\sqrt{z_j}\cos^i\sqrt{z_j}}{\sqrt{z_j}} - \dfrac{\sin\sqrt{z_1}\cos^i\sqrt{z_1}}{\sqrt{z_1}}\right] = \\ \qquad (\Delta t)^{-2}\left[\dfrac{\sin\sqrt{z_j}}{\sqrt{z_j}}\left(\frac{2}{3} - \frac{1}{6}\cos\sqrt{z_j}\right) - \dfrac{\sin\sqrt{z_1}}{\sqrt{z_1}}\left(\frac{2}{3} - \frac{1}{6}\cos\sqrt{z_1}\right)\right], \ j = 2,3. \end{cases} \tag{5.6}$$

Since $z_j = O(\Delta^2 t)$ both the coefficients and the right-hand sides in (5.6) are bounded as $\Delta t \to 0$. Hence, by solving (5.6) we find the weights (5.4) which define for all $(\mu_0, \nu_0) \in \mathfrak{R}$ a fourth-order discretization of $\partial/\partial x$ and $\partial/\partial y$.

The analogues of the Tables 4.1 and 4.2 are given by the Tables 5.1 and 5.2.

TABLE 5.1. Gain factors $\|\phi_c\|_\infty / \|\phi_e\|_\infty$ based on the weights (5.5)

(a) $\hat{\mu} = .25$, $\Delta\mu = 0$

$\hat{\nu} \setminus \Delta\nu$	5%	10%	20%	50%
.1	16603	4382	1229	301
.2	1200	323	95	29
.3	363	103	34	14
.4	224	67	24	12
.5	193	59	22	12

(b) $\hat{\nu} = .25$, $\Delta\nu = 0$

$\hat{\mu} \setminus \Delta\mu$	5%	10%	20%	50%
.1	233	60	16	3.4
.2	641	182	59	21
.3	2819	817	272	99
.4	59251	9242	1599	213
.5	5118	1367	389	89

Concerning the strategies, these tables show the same tendency as was observed in the case of an 8-point molecule. Comparing both type of molecules, the large gain factors listed in this section indicate that the phase error can be greatly reduced by employing a 16-point space-discretization molecule. Of course, in using a more sophisticated molecule, the effort to evaluate the semi-discrete approximation to the PDE will also increase. In general, however, molecules using many points seem to be more efficient than those based on only a few points.

It should be observed that Table 5.1(a) is identical to Table 4.1(a). This is due to the fact that $\delta(\mu)$ assumes only one value, viz. δ_0 (defined by (3.3)). Hence, the weights $\xi_j^{(l)}$ have, in this case, no influence on the value of $\|\phi\|_e$.

TABLE 5.2. Gain factors $\|\phi_c\|_\infty / \|\phi_e\|_\infty$ based on the weights (5.6)

(a) $\hat{\mu} = .25$, $\Delta\mu = 0$

$\hat{\nu} \setminus \Delta\nu$	5%	10%	20%	50%
.1	1418	758	443	344
.2	137	77	50	24
.3	35	21	15	11
.4	20	12	10	8
.5	16	10	8	7

(b) $\hat{\nu} = .25$, $\Delta\nu = 0$

$\hat{\mu} \setminus \Delta\mu$	5%	10%	20%	50%
.1	$2.0_{10}7$	$2.6_{10}6$	$3.4_{10}5$	$2.7_{10}4$
.2	$8.3_{10}5$	$1.2_{10}5$	$1.9_{10}4$	$2.6_{10}3$
.3	$2.7_{10}5$	$4.0_{10}4$	$6.8_{10}3$	$1.0_{10}3$
.4	$1.4_{10}5$	$2.1_{10}4$	$3.6_{10}3$	$5.3_{10}2$
.5	$8.4_{10}4$	$1.2_{10}4$	$2.2_{10}3$	$3.2_{10}2$

6. THE SHALLOW WATER EQUATIONS

Consider the basic, linearized form of the shallow water equations:

$$\frac{\partial \mathbf{u}}{\partial t} = -g\nabla h, \quad \frac{\partial h}{\partial t} = -h_0 \nabla \mathbf{u}, \quad t \in [0,T], \ (x,y) \in \mathbb{R}^2, \tag{6.1}$$

where $\mathbf{u}$ is the depth-averaged velocity, h is the depth below the moving water surface, h_0 is the depth when the water is in rest, and g is the acceleration of gravity. Initial and boundary conditions on the square $0 \leqslant x,y \leqslant L$ were taken from the exact solution

$$(\mathbf{u}, h - h_0) = \tfrac{1}{4}\sin((-\sqrt{2gh_0}\,t + x + y)\frac{2\pi}{L}) \cdot (1, 1, \sqrt{2h_0/g}). \tag{6.2}$$

This means that one period of the solution (in the space-direction) fits exactly on the square $[0,L]\times[0,L]$. This feature enables us to impose a periodicity condition on the boundaries, thus simulating the $\mathbb{R}^2$ as the space domain.

In our experiments we used $h_0 = 80$, $g = 10$, $L = 6_{10}5$ and $\Delta x = \Delta y = L/10$. For the fourth-order discretization operators $D_x = D_y$ we made several choices: (i) the classical line molecule, i.e. $\xi_1^{(0)} = 2/3$, $\xi_2^{(0)} = -1/12$, (ii) the adapted 8-point molecules (cf. Section 4) and (iii) the still more sophisticated 16-point molecules (cf. Section 5). Finally, for the time integration we used $\Delta t = 1800$.

To measure the dispersion of the various schemes, we integrated in time 10 periods of the solution and compared the position of the 20-th zero of $h - h_0$ with its analytical value, i.e. 150000. As, in general, the position of this 20-th zero does not coincide with a step point, a local refinement during the last integration step was performed to determine the correct zero of the numerical wave.

When written in the form (1.1) the eigenvalues of the matrix $\omega_x A + \omega_y B = 2\pi(A+B)/L$ are given by

$$\alpha_\pm = \pm\sqrt{gh_0(\omega_x^2 + \omega_y^2)} = \pm\frac{2\pi}{L}\sqrt{2gh_0}. \tag{6.3}$$

In the exponentially fitted method the fitting point (μ_0,ν_0) was chosen such that $\nu_0 = (\underline{\nu}+\overline{\nu})/2$. We observe that fitting at (μ_0,ν_0) automatically implies fitting at $(\mu_0,-\nu_0)$.

In addition to the linear system (6.1) we also integrated the *nonlinear* modifications which are closer to the actual shallow water equations:

$$\frac{\partial \mathbf{u}}{\partial t} = -g\nabla h - (\mathbf{u}\cdot\nabla)\mathbf{u}, \quad \frac{\partial h}{\partial t} = -h_0\nabla \mathbf{u}, \tag{6.4}$$

$$\frac{\partial \mathbf{u}}{\partial t} = -g\nabla h, \quad \frac{\partial h}{\partial t} = -\nabla(h\mathbf{u}), \tag{6.5}$$

$$\frac{\partial \mathbf{u}}{\partial t} = -g\nabla h - (\mathbf{u}\cdot\nabla)\mathbf{u}, \quad \frac{\partial h}{\partial t} = -\nabla(h\mathbf{u}). \tag{6.6}$$

The initial conditions are prescribed by (6.2) and we require again periodicity on the boundaries.

In Table 6.1 the time lag of $h - h_0$, defined by

$$t_{num}^{(20)} - t_{exact}^{(20)}$$

is listed for various choices of the frequency region $\mathcal{R}$ and the discretization weights $\{\xi_j^{(l)}\}$. Here, $t_{num}^{(20)}$ denotes the position of the 20-th zero (on the t-axis) and $t_{exact}^{(20)}$ is its analytical counterpart. For the non-linear equations (6.4)-(6.6), the values of $t_{exact}^{(20)}$ are not known analytically and have been determined previously, using an extremely small Δx and Δt value. These values are found to be 150 001.3, 149 995.7 and 149 960.0, respectively.

From this table we see that the results for the linear test equation (6.1) confirm the analysis of Section 3. By choosing the standard fourth-order space molecule and adapting only the RK scheme, a significant reduction of dispersion is obtained, even in the case where the uncertainty intervals are substantial. If, in addition, the space molecules are adapted to these intervals the dispersion is negligible, especially when 16-point molecules are employed. However, if the non-linearity of the equations increases, the gain factors decrease. Moreover, the benefit of using a 16-point molecule becomes less obvious. It should be observed that the scheme based on (5.5) behaves remarkebly better than the one that uses (5.6). This is explained by the fact that these problems have a solution with only one dominant frequency, whereas the underlying strategy of (5.6) is based on μ- and ν-intervals in which many frequencies are situated.

Concluding, the described techniques to reduce dispersion behave excellent in case of linear problems but loose part of their significance if nonlinearity increases.

TABLE 6.1 Time lag in $h-h_0$ due to dispersion, for conventional and adapted methods; $\Delta x = \Delta y = 60000$, $\Delta t = 1800$. For the linear problem (6.1), the exact values of μ and ν are respectively given by $\Delta x \omega_x = \pi/5 \simeq .6283$ and $\Delta t \alpha = \frac{6\pi}{25} \simeq .7540...$

$\mathcal{R}=[\underline{\mu},\bar{\mu}]\times[\underline{\nu},\bar{\nu}]$	$\{\xi_j^{(l)}\}$	(6.1)	(6.4)	(6.5)	(6.6)
0×0	(3.4)	1066	1018	960	929
$\frac{\pi}{5}\times\frac{6\pi}{25}$	(3.4)	0	-61	-146	-215
[.6,.64]×[.7,.8]	(3.4)	34	-30	-115	-184
[.6,.7]×[.7,.8]	(3.4)	-122	-186	-272	-344
[.6,.64]×[.7,.8]	(4.4)	29	-34	-114	-173
[.6,.64]×[.7,.8]	(4.5)	16	-25	-76	-97
[.6,.64]×[.7,.8]	(5.5)	3	-17	-36	-12
[.6,.64]×[.7,.8]	(5.6)	5	-58	-135	-183
[.6,.7]×[.7,.8]	(4.4)	-74	-136	-217	-276
[.6,.7]×[.7,.8]	(4.5)	14	-23	-70	-86
[.6,.7]×[.7,.8]	(5.5)	4	-25	-55	-44
[.6,.7]×[.7,.8]	(5.6)	5	-55	-128	-171

REFERENCES

[1] BRUSA, L. and L. NIGRO, *A one-step method for direct integration of structural dynamic equations,* Int. J. Num. Meth. in Engng. 15, 1980, 685-699.

[2] HOUWEN, P.J. VAN DER and B.P. SOMMEYER, *Linear multistep methods with reduced truncation error for periodic initial-value problems,* IMAJNA 4, 1984, 479-489.

[3] HOUWEN, P.J. VAN DER and B.P. SOMMEYER, *High order difference schemes with reduced dispersion for hyperbolic differential equations,* J. Comp. App. Math. 12 & 13, 1985, 145-161.

[4] STREHMEL, K. and R. WEINER, *Nichtlineare Stabilität und Phasenuntersuchung adaptiver Nyström-Runge-Kutta Methoden,* to appear in Computing.

[5] STREHMEL, K., *Stabilitäts- und Phasenuntersuchung von linear-impliziten Methoden für* $y'' = f(t,y)$, in: Proc. of the "4° Tagung Numerische Behandlung von gewöhnliche Differentialgleichungen" (ed. R. März), Berlin, 1984.

[6] THOMAS, R.M., *Phase properties of high order, almost P-stable formulae,* BIT 24, 1984, 225-238.