Available online at www.sciencedirect.com

ScienceDirect

Applied Numerical Mathematics 59 (2009) 677–692

APPLIED NUMERICAL MATHEMATICS

www.elsevier.com/locate/apnum

Unconditional stability of second-order ADI schemes applied to multi-dimensional diffusion equations with mixed derivative terms

K.J. in 't Hout [a,*], B.D. Welfert [b]

[a] *Department of Mathematics and Computer Science, University of Antwerp, Middelheimlaan 1, B-2020 Antwerp, Belgium*
[b] *Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287, USA*

Available online 21 March 2008

Abstract

We consider the unconditional stability of second-order ADI schemes in the numerical solution of finite difference discretizations of multi-dimensional diffusion problems containing mixed spatial-derivative terms. We investigate an ADI scheme proposed by Craig and Sneyd, an ADI scheme that is a modified version thereof, and an ADI scheme introduced by Hundsdorfer and Verwer. Both sufficient and necessary conditions are derived on the parameters of each of these schemes for unconditional stability in the presence of mixed derivative terms. Our main result is that, under a mild condition on its parameter θ, the second-order Hundsdorfer and Verwer scheme is unconditionally stable when applied to semi-discretized diffusion problems with mixed derivative terms in arbitrary spatial dimensions $k \geqslant 2$.

Keywords: Initial-boundary value problems; Diffusion equations; Numerical solution; Finite difference methods; ADI splitting schemes; von Neumann stability analysis; Fourier analysis

1. Introduction

The semi-discretization of initial-boundary value problems for multi-dimensional diffusion equations leads to initial value problems for huge systems of ordinary differential equations (ODEs),

$$U'(t) = F\big(t, U(t)\big) \quad (t \geqslant 0), \qquad U(0) = U_0, \tag{1.1}$$

with given vector-valued function F and initial vector U_0. As these initial value problems are stiff, implicit methods are desirable for their numerical solution. However, due to the very large sizes of the systems, standard implicit methods are ineffective, and moreover, often even unfeasible on present-day computers.

In this paper we study splitting schemes for the numerical solution of systems (1.1). We assume that the function F is decomposed into a sum

$$F(t, v) = F_0(t, v) + F_1(t, v) + \cdots + F_k(t, v) \tag{1.2}$$

of $k+1$ terms F_j (with $k \geqslant 2$) that are easier to handle than F itself. In our application, the term F_0 contains all contributions to F stemming from the mixed derivatives in the diffusion equation, and this term will always be treated

* Corresponding author.
E-mail addresses: karel.inthout@ua.ac.be (K.J. in 't Hout), welfert@asu.edu (B.D. Welfert).

doi:10.1016/j.apnum.2008.03.016

explicitly in the numerical time-integration. Next, for each $j \geqslant 1$, F_j represents the contribution to F stemming from the second-order derivative in the j-th spatial direction, and this term will always be treated *implicitly*.

1.1. Two general Alternating Direction Implicit schemes

Consider the initial value problem (1.1) with (1.2). Let time step $\Delta t > 0$. The first splitting scheme that we shall study defines an approximation $U_n \approx U(t_n)$, with $t_n = n\Delta t$, successively for $n = 1, 2, 3, \dots$ by

$$\begin{cases} Y_0 = U_{n-1} + \Delta t F(t_{n-1}, U_{n-1}), \\ Y_j = Y_{j-1} + \theta \Delta t (F_j(t_n, Y_j) - F_j(t_{n-1}, U_{n-1})), \quad j = 1, 2, \dots, k, \\ \widehat{Y}_0 = Y_0 + \sigma \Delta t (F_0(t_n, Y_k) - F_0(t_{n-1}, U_{n-1})), \\ \widetilde{Y}_0 = \widehat{Y}_0 + \mu \Delta t (F(t_n, Y_k) - F(t_{n-1}, U_{n-1})), \\ \widetilde{Y}_j = \widetilde{Y}_{j-1} + \theta \Delta t (F_j(t_n, \widetilde{Y}_j) - F_j(t_{n-1}, U_{n-1})), \quad j = 1, 2, \dots, k, \\ U_n = \widetilde{Y}_k. \end{cases} \tag{1.3}$$

Here $\theta > 0$, $\sigma > 0$ and μ denote real parameters that specify the scheme. It can be verified, through Taylor expansion, that (1.3) is of classical order[1] two if and only if $\sigma = \theta$, $\mu = \frac{1}{2} - \theta$ and θ arbitrary.

The first two lines of (1.3) constitute the so-called *Douglas scheme*, which defines Y_k as the approximation to $U(t)$ at the new time level $t = t_n$, cf. e.g. [8]. In the Douglas scheme, an explicit Euler predictor step is followed by k implicit but unidirectional corrector steps, whose purpose is to stabilize the predictor step. The well-known Alternating Direction Implicit (ADI) methods of Douglas and Rachford [4] and of Brian [1] and Douglas [3] are special instances of this scheme, with $F_0 = 0$ and $\theta = 1$, $\theta = \frac{1}{2}$, respectively. These special ADI methods were initially developed for application to the two- and three-dimensional heat equations. For a survey of their development, see e.g. Peaceman [9].

The Douglas scheme has classical order two provided $F_0 = 0$ and $\theta = \frac{1}{2}$. In our application to diffusion equations where mixed spatial-derivative terms are present, however, it always holds that $F_0 \neq 0$ and then the order of the Douglas scheme is just one. This reduction in order is a consequence of the fact that the F_0 term is treated in a simple, forward Euler fashion.

Craig and Sneyd [2] investigated the stability of the Douglas scheme for diffusion equations with mixed derivative terms and proved the positive result that to any given spatial dimension $k \geqslant 2$ there exists a value $\bar{\theta} = \bar{\theta}(k)$ such that the Douglas scheme is unconditionally stable whenever $\theta \geqslant \bar{\theta}$.

The purpose of the additional stages $\widehat{Y}_0$, $\widetilde{Y}_0$, $\widetilde{Y}_j$ in (1.3) is to increase the order to two again, for general $F_0 \neq 0$, while simultaneously retaining the favorable property of unconditional stability. If $\mu = 0$, then (1.3) stands for the so-called *Craig–Sneyd (CS) scheme*. It was proposed, in a slightly different form, in [2], where it was called the "iterated scheme". This scheme is of order two (for general F_0) if and only if $\sigma = \theta = \frac{1}{2}$. For spatial dimensions k equal to 2 or 3, the second-order CS scheme is indeed unconditionally stable when applied to diffusion equations with mixed derivative terms, but this stability property is unfortunately lost whenever $k \geqslant 4$, cf. [2] and Section 2.2. To arrive at more freedom, we have added an explicit stage $\widetilde{Y}_0$ to the CS scheme, involving the full right-hand side function F and the new parameter μ. We call the complete scheme (1.3) the *Modified Craig–Sneyd (MCS) scheme*.

The second splitting scheme that we shall study in this paper for (1.1), (1.2) defines an approximation $U_n \approx U(t_n)$, $n = 1, 2, 3, \dots$, by

$$\begin{cases} Y_0 = U_{n-1} + \Delta t F(t_{n-1}, U_{n-1}), \\ Y_j = Y_{j-1} + \theta \Delta t (F_j(t_n, Y_j) - F_j(t_{n-1}, U_{n-1})), \quad j = 1, 2, \dots, k, \\ \widetilde{Y}_0 = Y_0 + \mu \Delta t (F(t_n, Y_k) - F(t_{n-1}, U_{n-1})), \\ \widetilde{Y}_j = \widetilde{Y}_{j-1} + \theta \Delta t (F_j(t_n, \widetilde{Y}_j) - F_j(t_n, Y_k)), \quad j = 1, 2, \dots, k, \\ U_n = \widetilde{Y}_k \end{cases} \tag{1.4}$$

with real parameters $\theta > 0$ and μ. Clearly, this scheme can be viewed as an extension to the Douglas scheme as well. We refer to (1.4) as the *Hundsdorfer–Verwer (HV) scheme*. It has classical order two if and only if $\mu = \frac{1}{2}$ and θ arbitrary. For $\mu = \frac{1}{2}$ the scheme (1.4) is identical to a scheme formulated by Hundsdorfer [7, p. 222]. Further, in the

[1] I.e., the order for fixed non-stiff ODEs.

case of linear autonomous problems (1.1), it can be seen to be equivalent to an Approximate Matrix Factorization method discussed in Hundsdorfer and Verwer [8, p. 400].

The scheme (1.4) was developed [7,8] for the numerical solution of multi-dimensional convection–diffusion–reaction equations, arising e.g. in atmospheric chemistry. The application of (1.4) to equations containing mixed derivative terms has first been considered in in 't Hout and Welfert [6], for the case of two spatial dimensions.

1.2. Application to multi-dimensional diffusion equations

In this paper we investigate the two splitting schemes (1.3) and (1.4) for the numerical solution of semi-discretized k-dimensional diffusion equations

$$\frac{\partial u}{\partial t} = \sum_{i\neq j} d_{ij} u_{x_i x_j} + d_{11} u_{x_1 x_1} + d_{22} u_{x_2 x_2} + \cdots + d_{kk} u_{x_k x_k} \tag{1.5}$$

on the spatial domain $[0,1]^k$ with prescribed initial and boundary conditions. Here $D = [d_{ij}]_{1\leqslant i,j\leqslant k}$ is a given real matrix, which is always assumed to be symmetric positive semi-definite. In particular,

$$d_{jj} \geqslant 0 \quad \text{and} \quad |d_{ij}| \leqslant \sqrt{d_{ii} d_{jj}} \quad \text{for } 1 \leqslant i, j \leqslant k. \tag{1.6}$$

We are interested in the general case of (1.5) where actual mixed derivative terms are present and the spatial dimension $k \geqslant 2$ is arbitrary. Multi-dimensional diffusion problems of this type play an important role in a variety of applied areas, such as financial mathematics where they appear for instance in modeling the prices of options on a number of correlated assets, see e.g. Duffy [5] and Tavella and Randall [10].

For the semi-discretization of Eq. (1.5) we consider finite differences. All spatial derivatives in (1.5) are approximated using second-order central differences on a rectangular grid with constant mesh width $\Delta x_i > 0$ in the x_i-direction ($1 \leqslant i \leqslant k$):

$$(u_{x_i x_i})_\ell \approx \frac{u_{\ell+e_i} - 2u_\ell + u_{\ell-e_i}}{(\Delta x_i)^2}, \tag{1.7a}$$

$$\begin{aligned}(u_{x_i x_j})_\ell \approx{}& \frac{(1+\beta_{ij})(u_{\ell+e_i+e_j} + u_{\ell-e_i-e_j}) - (1-\beta_{ij})(u_{\ell-e_i+e_j} + u_{\ell+e_i-e_j})}{4\Delta x_i \Delta x_j} \\ &+ \frac{\beta_{ij}(4u_\ell - 2(u_{\ell+e_i} + u_{\ell+e_j} + u_{\ell-e_i} + u_{\ell-e_j}))}{4\Delta x_i \Delta x_j}, \quad i \neq j.\end{aligned} \tag{1.7b}$$

Here $\ell = (\ell_1, \ell_2, \ldots, \ell_k)$ and the unit vectors $e_1, e_2, \ldots, e_k$ denote multi-indices, u_ℓ stands for $u(\ell_1\Delta x_1, \ell_2\Delta x_2, \ldots, \ell_k\Delta x_k, t)$, and β_{ij} are real parameters with $\beta_{ij} = \beta_{ji}$. The right-hand side of (1.7b) is the most general formula of a second-order finite difference approximation to the cross derivative $u_{x_i x_j}$ based on a centered 9-point stencil. When $\beta_{ij} \equiv 0$, (1.7b) reduces to the standard 4-point formula

$$(u_{x_i x_j})_\ell \approx \frac{u_{\ell+e_i+e_j} + u_{\ell-e_i-e_j} - u_{\ell-e_i+e_j} - u_{\ell+e_i-e_j}}{4\Delta x_i \Delta x_j}.$$

With the specific arrangement of terms in Eq. (1.5), semi-discretization naturally leads to a partitioned system (1.1), (1.2) with $F_j(t,v) = A_j(t)v + g_j(t)$ ($0 \leqslant j \leqslant k$), where $A_0(t)$ is a matrix that represents all mixed derivative terms in (1.5) and, for each $j \geqslant 1$, $A_j(t)$ is a matrix that represents the second-order spatial derivative in the x_j-direction. The functions g_j are obtained from the boundary condition. The entries of the solution vector $U(t)$ to (1.1) constitute approximations of the exact solution values $u(x,t)$ to (1.5) at the spatial grid points, ordered in a convenient way.

1.3. Stability analysis of (1.3), (1.4)

The purpose of this paper is to assess the stability of the schemes (1.3) and (1.4) in the application to finite difference discretizations (1.7) of initial-boundary value problems for the k-dimensional diffusion equation (1.5). Here we are interested in *unconditional* stability, i.e., without any restriction on the time step $\Delta t > 0$.

Our analysis is equivalent to the well-known von Neumann (Fourier) analysis. Accordingly, stability is always considered in the l_2-norm and, in order to make the analysis feasible, all coefficients d_{ij} in (1.5) are assumed to be constant and the boundary condition for (1.5) to be periodic. Under these assumptions, the matrices $A_0, A_1, \ldots, A_k$ obtained by finite difference discretization are constant and form Kronecker products of circulant matrices. Hence, they are normal and commute with each other. This implies that stability can be analyzed by considering the linear scalar ODE

$$U'(t) = (\lambda_0 + \lambda_1 + \cdots + \lambda_k)U(t), \tag{1.8}$$

where λ_j denotes an eigenvalue of the matrix A_j, $0 \leqslant j \leqslant k$.

When applied to (1.8), the schemes (1.3), (1.4) reduce to the scalar iterations

$$U_n = S(z_0, z_1, \ldots, z_k)U_{n-1} \tag{1.9}$$

and

$$U_n = T(z_0, z_1, \ldots, z_k)U_{n-1}, \tag{1.10}$$

respectively, with $z_j = \lambda_j \Delta t$ $(0 \leqslant j \leqslant k)$ and

$$S(z_0, z_1, \ldots, z_k) = 1 + \frac{z_0 + z}{p} + \sigma \frac{z_0(z_0 + z)}{p^2} + \mu \frac{(z_0 + z)^2}{p^2}, \tag{1.11}$$

$$T(z_0, z_1, \ldots, z_k) = 1 + 2\frac{z_0 + z}{p} - \frac{z_0 + z}{p^2} + \mu \frac{(z_0 + z)^2}{p^2}, \tag{1.12}$$

where we have used the notation

$$z = z_1 + z_2 + \cdots + z_k \quad \text{and} \quad p = (1 - \theta z_1)(1 - \theta z_2) \cdots (1 - \theta z_k). \tag{1.13}$$

It is readily verified, upon substituting discrete Fourier modes into (1.7), that the scaled eigenvalues z_j are given by

$$z_0 = \sum_{i \neq j} r_{ij} d_{ij} \bigl[-\sin\phi_i \sin\phi_j + \beta_{ij}(1 - \cos\phi_i)(1 - \cos\phi_j)\bigr], \tag{1.14a}$$

$$z_j = -2r_{jj} d_{jj}(1 - \cos\phi_j), \quad j = 1, 2, \ldots, k, \tag{1.14b}$$

where the angles ϕ_j are integer multiples of $2\pi/m_j$ with m_j the dimension of the grid in the x_j-direction $(1 \leqslant j \leqslant k)$. Further,

$$r_{ij} = \frac{\Delta t}{\Delta x_i \Delta x_j} \quad \text{for } 1 \leqslant i, j \leqslant k.$$

The scheme (1.3) is unconditionally stable when applied to the finite difference discretization of the initial-boundary value problem for (1.5) as described above if and only if

$$\bigl|S(z_0, z_1, \ldots, z_k)\bigr| \leqslant 1 \tag{1.15}$$

for all $(z_0, z_1, \ldots, z_k)$ given by (1.14) and $\Delta t > 0$. For the scheme (1.4) the same result holds, provided that condition (1.15) is replaced by

$$\bigl|T(z_0, z_1, \ldots, z_k)\bigr| \leqslant 1. \tag{1.16}$$

In the special case where all mixed derivative terms in (1.5) are absent, conclusions about the unconditional stability of (1.3), (1.4) are readily obtained. It can be shown, using (1.15) or (1.16) with $z_0 = 0$, that in this case the CS scheme is unconditionally stable for $\theta \geqslant \frac{1}{2}$, the second-order MCS scheme for $\theta \geqslant \frac{1}{4}$, and the second-order HV scheme also for $\theta \geqslant \frac{1}{4}$. Note that these parameter ranges are all independent of the spatial dimension k.

The analysis of unconditional stability of the schemes (1.3), (1.4) in the general case of (1.5), with mixed derivative terms, is essentially more complicated, and this forms the objective of our paper.

1.4. Outline of this paper

Section 2.1 contains preliminary results on the eigenvalues z_j given by (1.14).

In Section 2.2 we study the stability of the CS scheme. We prove a sufficient condition for $k = 2$ and $k = 3$ and a necessary condition for arbitrary $k \geqslant 2$ on the parameters θ, σ of the CS scheme for unconditional stability in the application to semi-discrete k-dimensional diffusion problems (1.5) with mixed derivative terms. In particular, these results imply that the second-order CS scheme ($\sigma = \theta = \frac{1}{2}$) is unconditionally stable if the spatial dimension equals 2 or 3, but not, in general, in any larger dimension.

In Section 2.3 we investigate the MCS scheme. We assume that $\sigma = \theta$ and $\mu = \frac{1}{2} - \theta$, so that the scheme is of order 2. As in Section 2.2, we arrive at a useful sufficient condition on the parameter θ for unconditional stability in two and three spatial dimensions, and a necessary condition relevant to arbitrary spatial dimensions. It is conjectured that this necessary condition is also sufficient.

In Section 2.4 we consider the HV scheme and present the main result of our paper, Theorem 2.8. This theorem provides a practical condition on the parameter θ which guarantees that the second-order HV scheme is unconditionally stable in the application to diffusion problems (1.5) with mixed derivative terms in *arbitrary* spatial dimensions $k \geqslant 2$. It is subsequently shown that this sufficient condition on θ is also necessary.

Section 2.5 summarizes and compares the unconditional stability results for the ADI schemes obtained in Sections 2.2, 2.3, 2.4.

Appendix A is devoted to the proof of Theorem 2.8.

2. Unconditional stability of (1.3) and (1.4)

To render our analysis feasible, we will always make the minor assumption that there exist real numbers $\delta_1, \delta_2, \ldots, \delta_k$ such that the parameters β_{ij} in (1.7b) satisfy

$$\beta_{ij} = -\delta_i \delta_j, \quad |\delta_j| \leqslant 1, \ 1 \leqslant i \neq j \leqslant k. \tag{2.1}$$

Further we recall the notation (1.13), which is used throughout this paper.

2.1. Preliminaries

Our subsequent analysis is based upon four properties of the scaled eigenvalues z_j given by (1.14). These key properties are formulated in

Lemma 2.1. *Let* $z_0, z_1, \ldots, z_k$ *satisfy* (1.14). *Then*:

$$\text{all } z_j \text{ are real}, \tag{2.2a}$$

$$z_j \leqslant 0 \quad \text{for } 1 \leqslant j \leqslant k, \tag{2.2b}$$

$$z + z_0 \leqslant 0, \tag{2.2c}$$

$$|z_0| \leqslant \sum_{i \neq j} \sqrt{z_i z_j}. \tag{2.2d}$$

Proof. The properties (2.2a), (2.2b) are obvious. We consider (2.2c) and (2.2d). For $j = 1, 2, \ldots, k$ define

$$\mathbf{v}_j = \begin{bmatrix} \sin\phi_j \\ \delta_j(1-\cos\phi_j) \end{bmatrix} \quad \text{and} \quad x_j = \sqrt{2 r_{jj} d_{jj}(1-\cos\phi_j)}.$$

Then, using (2.1), we have

$$z_0 = -\sum_{i \neq j} r_{ij} d_{ij} \mathbf{v}_i \cdot \mathbf{v}_j \quad \text{and} \quad z_j = -x_j^2 \quad (1 \leqslant j \leqslant k). \tag{2.3}$$

Next,

$$r_{jj} d_{jj} \|\mathbf{v}_j\|_2^2 \leqslant r_{jj} d_{jj} \{\sin^2\phi_j + (1-\cos\phi_j)^2\} = x_j^2. \tag{2.4}$$

From (2.3), (2.4), $r_{ij} = \sqrt{r_{ii} r_{jj}}$ and the non-negativity of D, we obtain

$$\begin{aligned}
-(z + z_0) &= \sum_{j=1}^{k} x_j^2 + \sum_{i \neq j} r_{ij} d_{ij} \mathbf{v}_i \cdot \mathbf{v}_j \\
&\geqslant \sum_{i,j=1}^{k} r_{ij} d_{ij} \mathbf{v}_i \cdot \mathbf{v}_j \\
&= \mathbf{u}^T D \mathbf{u} + \mathbf{w}^T D \mathbf{w} \\
&\geqslant 0,
\end{aligned}$$

where

$$\mathbf{u} = [\sqrt{r_{jj}} \sin \phi_j]_{1 \leqslant j \leqslant k} \quad \text{and} \quad \mathbf{w} = \left[\sqrt{r_{jj}} \delta_j (1 - \cos \phi_j)\right]_{1 \leqslant j \leqslant k}.$$

Subsequently, using (1.6), the Cauchy–Schwarz inequality and (2.4), we arrive at the bound

$$\begin{aligned}
|z_0| &\leqslant \sum_{i \neq j} r_{ij} |d_{ij}| |\mathbf{v}_i \cdot \mathbf{v}_j| \\
&\leqslant \sum_{i \neq j} \sqrt{r_{ii} r_{jj}} \sqrt{d_{ii} d_{jj}} \|\mathbf{v}_i\|_2 \|\mathbf{v}_j\|_2 \\
&\leqslant \sum_{i \neq j} x_i x_j \\
&= \sum_{i \neq j} \sqrt{z_i z_j},
\end{aligned}$$

which concludes the proof. □

It follows from (2.2a,b) that

$$p \geqslant 1 - \theta z \geqslant 1. \tag{2.5}$$

Next, as a corollary to (2.2a,b,d), we have

$$z + z_0 \geqslant z - |z_0| \geqslant \sum_{j=1}^{k} z_j - \sum_{i \neq j} \sqrt{z_i z_j} = -\left(\sum_{j=1}^{k} \sqrt{-z_j} \right)^2. \tag{2.6}$$

If $k = 2$, then the bound (2.2d) reduces to $|z_0| \leqslant 2\sqrt{z_1 z_2}$. It is interesting to remark that in [6] a more general bound was considered when $k = 2$, for *complex* numbers z_0, z_1, z_2:

$$|z_0| \leqslant 2\sqrt{\Re z_1 \Re z_2}.$$

In [6] positive results were obtained on the question of whether the latter bound, together with $\Re z_1, \Re z_2 \leqslant 0$, implies the stability conditions (1.15) and (1.16) for $k = 2$.

In the rest of this paper we will always assume that (2.2) holds. We shall frequently employ the notation

$$y_j = \sqrt{-\theta z_j} \quad (1 \leqslant j \leqslant k). \tag{2.7}$$

Then

$$p = \prod_{j=1}^{k} \left(1 + y_j^2\right) \quad \text{and} \quad z = -\frac{1}{\theta} \sum_{j=1}^{k} y_j^2. \tag{2.8}$$

Further, (2.6) becomes

$$z + z_0 \geqslant -\frac{1}{\theta} \left(\sum_{j=1}^{k} y_j \right)^2. \tag{2.9}$$

2.2. Unconditional stability of the Craig–Sneyd scheme

In this section we study the stability of the CS scheme, i.e. (1.3) with $\mu = 0$. It is easily verified that the stability requirement (1.15), for $\mu = 0$, is equivalent to

$$2p^2 + pz + (p + \sigma z)z_0 + \sigma z_0^2 \geqslant 0 \quad \text{and} \quad p + \sigma z_0 \geqslant 0. \tag{2.10}$$

We are interested in the case where the CS scheme has order 2. Recall from Section 1.1 that this holds if and only if $\sigma = \theta = \frac{1}{2}$. The following result guarantees unconditional stability of the second-order CS scheme when it is applied to semi-discretized diffusion equations (1.5) in two and three spatial dimensions.

Theorem 2.2. *Let $k = 2$ or $k = 3$. Consider Eq.* (1.5) *with positive semi-definite matrix D and periodic boundary condition. Assume* (1.1), (1.2) *is obtained after semi-discretization and splitting of* (1.5) *as described in Section* 1.2. *Let $\sigma = \frac{1}{2}$ and $\mu = 0$. Then the scheme* (1.3) *is unconditionally stable when applied to* (1.1), (1.2) *whenever $\theta \geqslant \frac{1}{2}$.*

We remark that the above result was already stated in [2] when $\beta_{ij} \equiv 0$. The proof in loc. cit. relied in part upon numerical evidence, however. Below a complete, short proof of Theorem 2.2 is given.

Proof. Let z_j $(0 \leqslant j \leqslant k)$ satisfy (1.14) with $k = 2$ or $k = 3$. We will prove that (2.10) holds, which yields the result. For $\sigma = \frac{1}{2}$ the condition (2.10) reduces to

$$p + \frac{1}{2}z_0 \geqslant 0,$$

in view of $2p^2 + pz + (p + \sigma z)z_0 + \sigma z_0^2 = (p + \frac{1}{2}z_0)(2p + z) + \frac{1}{2}z_0^2$ and (2.5). Using (2.7)–(2.9), we find that this inequality is fulfilled if

$$P(y_1, y_2, \ldots, y_k) := \prod_{j=1}^{k}\left(1 + y_j^2\right) + \frac{1}{2\theta}\sum_{j=1}^{k} y_j^2 - \frac{1}{2\theta}\left(\sum_{j=1}^{k} y_j\right)^2 \geqslant 0$$

whenever $y_j \geqslant 0$ $(1 \leqslant j \leqslant k)$. For $k = 2$, we obtain (use that $\theta \geqslant \frac{1}{2}$)

$$P(y_1, y_2) \geqslant \left(1 + y_1^2\right)\left(1 + y_2^2\right) - 2y_1y_2 = 1 + (y_1 - y_2)^2 + y_1^2y_2^2 \geqslant 1.$$

Next, for $k = 3$ we have

$$\begin{aligned} P(y_1, y_2, y_3) &\geqslant 1 + \sum_j y_j^2 + \sum_{i<j}\left(y_i^2y_j^2 - 2y_iy_j\right) + y_1^2y_2^2y_3^2 \\ &= 1 + \frac{1}{2}\sum_{i<j}(y_i - y_j)^2 + \sum_{i<j}\left(y_i^2y_j^2 - y_iy_j\right) + y_1^2y_2^2y_3^2 \\ &= \frac{1}{4} + \frac{1}{2}\sum_{i<j}(y_i - y_j)^2 + \sum_{i<j}\left(y_iy_j - \frac{1}{2}\right)^2 + y_1^2y_2^2y_3^2 \\ &\geqslant \frac{1}{4}. \qquad \Box \end{aligned}$$

The above proof does not extend to dimensions $k \geqslant 4$. For example, if $y = \frac{1}{2}\sqrt{2}$, then $P(y, y, y, y, 0, \ldots, 0) = -\frac{15}{16} < 0$.

The subsequent result yields a necessary condition on the parameters of the general CS scheme for unconditional stability when applied to diffusion equations (1.5) in arbitrary spatial dimensions k.

Theorem 2.3. *Let $k \geqslant 2$. Suppose that the scheme* (1.3) *with $\mu = 0$ is unconditionally stable whenever it is applied to a system* (1.1), (1.2) *that is obtained by semi-discretization and splitting as described in Section* 1.2 *of any Eq.* (1.5) *with positive semi-definite $k \times k$ matrix D and periodic boundary condition. Then θ must satisfy the bound*

$$\theta \geqslant \max\left\{\frac{1}{2}, \sigma c_k k\right\} \quad \text{with } c_k = \left(1 - \frac{1}{k}\right)^k. \tag{2.11}$$

In particular, $\theta > \sigma$ whenever $k \geqslant 4$.

Proof. Consider the $k \times k$ matrix $D = [d_{ij}]$ with $d_{ij} \equiv 1$. Clearly, D is symmetric and positive semi-definite. Let $\Delta x_i \equiv \Delta x > 0$, so that $r_{ij} \equiv r := \Delta t/(\Delta x)^2$.

First, choose the angles ϕ_j in (1.14) equal to zero for $j \geqslant 2$. Then we obtain eigenvalues z_j given by $z_0 = 0$, $z_1 = -2r(1 - \cos\phi_1)$, $z_2 = z_3 = \cdots = z_k = 0$. The assumption in the theorem and the stability criterion (2.10) imply that

$$2p + z = 2 + (1 - 2\theta)z_1 \geqslant 0$$

whenever $r > 0$, $\phi_1 \in [0, 2\pi)$. This immediately implies $\theta \geqslant \frac{1}{2}$.

Next, choose all angles ϕ_j in (1.14) the same, i.e., $\phi_j \equiv \phi$. We then have eigenvalues z_j given by

$$z_j = -2r(1 - \cos\phi), \quad j = 1, 2, \ldots, k, \tag{2.12a}$$

$$z_0 = -rk(k-1)\left[\sin^2\phi - \bar{\beta}(1 - \cos\phi)^2\right], \tag{2.12b}$$

where

$$\bar{\beta} = \frac{\sum_{i \neq j} \beta_{ij}}{k(k-1)}. \tag{2.13}$$

Note that $|\bar{\beta}| \leqslant 1$. The assumption in the theorem and the stability criterion (2.10) imply that $p + \sigma z_0 \geqslant 0$ whenever $z_0, z_1, \ldots, z_k$ are given by (2.12) with arbitrary $r > 0$, $\phi \in [0, 2\pi)$. As a consequence,

$$\theta \geqslant \sigma\left(\frac{-\theta z_0}{p}\right) = \sigma k(k-1)\frac{\theta r[\sin^2\phi - \bar{\beta}(1 - \cos\phi)^2]}{[1 + 2\theta r(1 - \cos\phi)]^k}$$

whenever $r > 0$, $\phi \in [0, 2\pi)$. Setting $\alpha = 2\theta r(1 - \cos\phi)$, we get

$$\theta \geqslant \sigma k(k-1)\frac{\alpha[1 + \cos\phi - \bar{\beta}(1 - \cos\phi)]/2}{(1 + \alpha)^k}$$

for all $\alpha > 0$, $\phi \in (0, 2\pi)$. Taking the supremum over ϕ, gives

$$\theta \geqslant \sigma k(k-1)\frac{\alpha}{(1 + \alpha)^k} \quad \text{for all } \alpha > 0.$$

It is easily seen that the right-hand side is maximal when $\alpha = 1/(k-1)$. Upon inserting this value, the bound (2.11) is obtained. Finally, it is readily verified that $c_k k > 1$ for $k \geqslant 4$, which proves the last statement of the theorem. □

An important corollary to Theorem 2.3 is that, for any fixed $k \geqslant 4$, if the CS scheme is required to be unconditionally stable in the application to standard finite difference discretizations of the k-dimensional diffusion equation (1.5), with mixed derivative terms, then $\theta > \sigma$ and hence its order is at most one.

It is interesting to remark that if $k \geqslant 4$ the necessary condition (2.11) combined with $\sigma \geqslant \frac{1}{2}$ is identical to the *sufficient* condition that was given in [2] for unconditional stability of the CS scheme. But the derivation in loc. cit. was based in part upon numerical evidence. A rigorous proof of the (conjectured) sufficiency of this condition when $k \geqslant 4$ is still lacking at present.[2]

Section 2.5 provides numerical values of the bound (2.11) for $\sigma = \frac{1}{2}$ and $2 \leqslant k \leqslant 9$.

[2] Such a proof appears to be more complicated than a proof given further on in this paper for the HV scheme.

2.3. Unconditional stability of the Modified Craig–Sneyd scheme

In this section we study the MCS scheme (1.3) with $\sigma=\theta$, $\mu=\frac{1}{2}-\theta$. Recall that this scheme is of order 2 for any given θ.

For arbitrary parameters θ, σ, μ the stability requirement (1.15) is equivalent to (2.14), where

$$2p^2+pz+\mu z^2+\big[p+(\sigma+2\mu)z\big]z_0+(\sigma+\mu)z_0^2\geqslant 0, \tag{2.14a}$$

$$p-\sigma z+(\sigma+\mu)(z_0+z)\geqslant 0. \tag{2.14b}$$

A preliminary result is

Lemma 2.4. *If $\sigma=\theta$ and $\mu=\frac{1}{2}-\theta$, then* (2.14a) *holds.*

Proof. Condition (2.14a) is fulfilled if the discriminant

$$\begin{aligned}\Delta&=\big[p+(1-\theta)z\big]^2-2\bigg[2p^2+pz+\left(\frac{1}{2}-\theta\right)z^2\bigg]\\&=-3p^2-2\theta pz+\theta^2z^2\\&=-(p+\theta z)(3p-\theta z)\end{aligned}$$

is non-positive, and this is a direct consequence of (2.5). □

For the case of two- and three-dimensional diffusion equations, we have the following positive result guaranteeing unconditional stability of the second-order MCS scheme.

Theorem 2.5. *Let $k=2$ or $k=3$. Consider Eq.* (1.5) *with positive semi-definite matrix D and periodic boundary condition. Assume* (1.1), (1.2) *is obtained after semi-discretization and splitting of* (1.5) *as described in Section* 1.2. *Let $\sigma=\theta$ and $\mu=\frac{1}{2}-\theta$. Then the scheme* (1.3) *is unconditionally stable when applied to* (1.1), (1.2) *whenever $\theta\geqslant\frac{1}{3}$ (if $k=2$) or $\theta\geqslant\frac{6}{13}$ (if $k=3$).*

Proof. In view of Lemma 2.4 it remains to show that condition (2.14b) is fulfilled for the eigenvalues z_j given by (1.14) with $k=2$ or $k=3$. Using (2.7)–(2.9), we obtain that (2.14b) holds if

$$\theta\geqslant\frac{1}{2}\frac{(\sum_{j=1}^k y_j)^2}{\prod_{j=1}^k(1+y_j^2)+\sum_{j=1}^k y_j^2} \tag{2.15}$$

whenever $y_j\geqslant 0$ ($1\leqslant j\leqslant k$).

For $k=2$, the right-hand side of (2.15) is bounded from above by $\frac{1}{3}$ if

$$2\big[\big(1+y_1^2\big)\big(1+y_2^2\big)+y_1^2+y_2^2\big]\geqslant 3(y_1+y_2)^2.$$

Now

$$\begin{aligned}2\big[\big(1+y_1^2\big)\big(1+y_2^2\big)+y_1^2+y_2^2\big]-3(y_1+y_2)^2&=2+y_1^2+y_2^2+2y_1^2y_2^2-6y_1y_2\\&=(y_1-y_2)^2+2(y_1y_2-1)^2\geqslant 0.\end{aligned}$$

For $k=3$, the right-hand side of (2.15) is bounded from above by $\frac{6}{13}$ if

$$12\big[\big(1+y_1^2\big)\big(1+y_2^2\big)\big(1+y_3^2\big)+y_1^2+y_2^2+y_3^2\big]\geqslant 13(y_1+y_2+y_3)^2.$$

Rewriting this inequality yields

$$11\big(y_1^2+y_2^2+y_3^2\big)+12\big(1+y_1^2y_2^2+y_1^2y_3^2+y_2^2y_3^2+y_1^2y_2^2y_3^2\big)\geqslant 26(y_1y_2+y_1y_3+y_2y_3),$$

and subsequently

$$\frac{11}{2}\sum_{i<j}(y_i-y_j)^2+12\big(1+y_1^2y_2^2+y_1^2y_3^2+y_2^2y_3^2+y_1^2y_2^2y_3^2\big)\geqslant 15(y_1y_2+y_1y_3+y_2y_3).$$

Put $u = y_1y_2$, $v = y_1y_3$, $w = y_2y_3$ and consider

$$P(u,v,w) = 4\bigl(1 + u^2 + v^2 + w^2 + uvw\bigr) - 5(u + v + w).$$

If $P(u,v,w) \geqslant 0$ whenever $u,v,w \geqslant 0$, then the proof for $k = 3$ is complete. It is easily verified that $P(u,v,w) > 0$ whenever $u = 0$ or $v = 0$ or $w = 0$. Next, it is easily seen that $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is the only critical point of P in the region $u,v,w > 0$ and the value of P at this point equals 0. Finally, $P(u,v,w) \to \infty$ as $(u,v,w) \to \infty$. Consequently, $P(u,v,w) \geqslant 0$ whenever $u,v,w \geqslant 0$. □

The subsequent result gives a necessary condition on the parameter θ, defining the second-order MCS scheme, for unconditional stability in the case of diffusion equations (1.5) in arbitrary spatial dimensions k.

Theorem 2.6. *Let $k \geqslant 2$. Suppose that the scheme* (1.3) *with $\sigma = \theta$, $\mu = \frac{1}{2} - \theta$ is unconditionally stable whenever it is applied to a system* (1.1), (1.2) *that is obtained by semi-discretization and splitting as described in Section* 1.2 *of any Eq.* (1.5) *with positive semi-definite $k \times k$ matrix D and periodic boundary condition. Then θ satisfies the bound*

$$\theta \geqslant \frac{1}{2} b_k k \quad \text{with } b_k = \frac{1}{1 + (1 + 1/(k-1))^{k-1}}. \tag{2.16}$$

Proof. The assumption in the theorem and the stability criterion (2.14) imply that

$$p - \theta z + \frac{1}{2}(z_0 + z) \geqslant 0$$

whenever $z_0, z_1, \ldots, z_k$ are given by (2.12) with arbitrary $r > 0$, $\phi \in [0, 2\pi)$. Thus, with $\bar{\beta}$ defined by (2.13), we have

$$\theta \geqslant -\frac{1}{2}\frac{\theta z_0 + \theta z}{p - \theta z} = \frac{k}{2}\frac{(k-1)\theta r[\sin^2\phi - \bar{\beta}(1 - \cos\phi)^2] + 2\theta r(1 - \cos\phi)}{[1 + 2\theta r(1 - \cos\phi)]^k + 2k\theta r(1 - \cos\phi)}$$

whenever $r > 0$, $\phi \in [0, 2\pi)$. Setting $\alpha = 2\theta r(1 - \cos\phi)$, as before, we obtain

$$\theta \geqslant \frac{k}{2}\frac{(k-1)\alpha[1 + \cos\phi - \bar{\beta}(1 - \cos\phi)]/2 + \alpha}{(1 + \alpha)^k + k\alpha}$$

for all $\alpha > 0$, $\phi \in (0, 2\pi)$. Taking the supremum over ϕ, gives

$$\theta \geqslant \frac{k}{2}\frac{k\alpha}{(1 + \alpha)^k + k\alpha} \quad \text{for all } \alpha > 0.$$

The right-hand side is maximal when $\alpha = 1/(k-1)$, and upon inserting this value (2.16) is obtained. □

It is easily verified that (2.16) reduces to the sufficient conditions on θ given in Theorem 2.5 for $k = 2$ and $k = 3$. Hence, these conditions are optimal.

We conjecture that for any given dimension $k \geqslant 2$ the bound (2.16) is sufficient for unconditional stability of the second-order MCS scheme when applied to semi-discretized diffusion equations with mixed derivative terms.[3]

Section 2.5 provides numerical values of the bound (2.16) for $2 \leqslant k \leqslant 9$.

2.4. Unconditional stability of the Hundsdorfer–Verwer scheme

We now consider the HV scheme (1.4). Recall that this scheme has order 2 if $\mu = \frac{1}{2}$, independent of θ.

For arbitrary parameters θ, μ the stability requirement (1.16) is equivalent to (2.17), where

$$2p^2 + (2p - 1)(z_0 + z) + \mu(z_0 + z)^2 \geqslant 0, \tag{2.17a}$$

$$2p - 1 + \mu(z_0 + z) \geqslant 0. \tag{2.17b}$$

First we have

[3] As for the original CS scheme, a proof of this conjecture appears to be more complicated than the proof given further on in this paper for the HV scheme.

Lemma 2.7. *If $\mu \geqslant \frac{1}{2}$, then* (2.17a) *holds.*

Proof. Condition (2.17a) holds if the discriminant $\Delta = (2p-1)^2 - 8\mu p^2$ is non-positive and this is fulfilled if $\mu \geqslant \frac{1}{2}$ (use that $p \geqslant 1$). □

The following theorem constitutes the main result of our paper. It guarantees unconditional stability of the second-order HV scheme when applied to semi-discretized diffusion equations with mixed derivative terms in arbitrary spatial dimensions k.

Theorem 2.8. *Let $k \geqslant 2$. Consider Eq.* (1.5) *with positive semi-definite $k \times k$ matrix D and periodic boundary condition. Assume* (1.1), (1.2) *is obtained after semi-discretization and splitting of* (1.5) *as described in Section* 1.2. *Let $\mu \geqslant \frac{1}{2}$. Then the scheme* (1.4) *is unconditionally stable when applied to* (1.1), (1.2) *whenever*

$$\theta \geqslant \mu a_k k, \tag{2.18}$$

where a_k is the unique solution $a \in (0, \frac{1}{2})$ of

$$2a\left(1 + \frac{1-a}{k-1}\right)^{k-1} - 1 = 0. \tag{2.19}$$

Proof. In view of Lemma 2.7 it remains to prove the inequality (2.17b) for the eigenvalues z_j given by (1.14). Using (2.7)–(2.9), we find that this holds if

$$\theta \geqslant \mu \frac{(\sum_{j=1}^{k} y_j)^2}{2\prod_{j=1}^{k}(1+y_j^2) - 1} \tag{2.20}$$

whenever $y_j \geqslant 0$ $(1 \leqslant j \leqslant k)$. It is proved in Appendix A that the right-hand side of (2.20) attains its maximum when $y_1 = y_2 = \cdots = y_k \in [0, 1]$. Hence, (2.20) is fulfilled whenever $y_j \geqslant 0$ $(1 \leqslant j \leqslant k)$ if

$$\theta \geqslant \mu \max_{0 \leqslant \alpha \leqslant 1} \frac{k^2 \alpha}{2(1+\alpha)^k - 1}. \tag{2.21}$$

Subsequently, in Appendix A it is shown that (2.21) is equivalent to (2.18). □

Theorem 2.8 appears to be the first result in the literature on ADI schemes that guarantees both second order ($\mu = \frac{1}{2}$) and unconditional stability when applied to diffusion equations with mixed derivative terms in *arbitrary* spatial dimension k. Up to now, this was only (essentially) known to be possible in two and three spatial dimensions, cf. [2] and Theorem 2.2.

We note that for the two-dimensional case, $k = 2$, the result of Theorem 2.8 was previously obtained in [6, Theorem 3.2].

The next result shows that the sufficient condition (2.18) is optimal, for any given dimension k.

Theorem 2.9. *Let $k \geqslant 2$. Suppose that the scheme* (1.4) *is unconditionally stable whenever it is applied to a system* (1.1), (1.2) *that is obtained by semi-discretization and splitting as described in Section* 1.2 *of any Eq.* (1.5) *with positive semi-definite $k \times k$ matrix D and periodic boundary condition. Then θ satisfies the bound* (2.18).

Proof. The assumption in the theorem and the stability criterion (2.17) imply that

$$2p - 1 + \mu(z_0 + z) \geqslant 0$$

whenever $z_0, z_1, \ldots, z_k$ are given by (2.12) with arbitrary $r > 0$, $\phi \in [0, 2\pi)$. It follows that, with $\bar{\beta}$ defined by (2.13),

$$\theta \geqslant -\mu \frac{\theta z_0 + \theta z}{2p-1} = \mu k \frac{(k-1)\theta r[\sin^2\phi - \bar{\beta}(1-\cos\phi)^2] + 2\theta r(1-\cos\phi)}{2[1 + 2\theta r(1-\cos\phi)]^k - 1}$$

whenever $r > 0$, $\phi \in [0, 2\pi)$. Setting $\alpha = 2\theta r(1-\cos\phi)$, this yields

$$\theta \geqslant \mu k \frac{(k-1)\alpha[1 + \cos\phi - \bar{\beta}(1-\cos\phi)]/2 + \alpha}{2(1+\alpha)^k - 1}$$

Table 1
Lower stability bounds on θ vs. $k = 2, 3, \ldots, 9$ for the CS scheme with $\sigma = \frac{1}{2}$, the second-order MCS scheme, and the second-order HV scheme

k	2	3	4	5	6	7	8	9
(2.11)	0.500	0.500	0.633	0.819	1.005	1.190	1.374	1.559
(2.16)	0.333	0.462	0.593	0.726	0.860	0.994	1.128	1.262
(2.18)	0.293	0.402	0.515	0.630	0.745	0.860	0.975	1.091

for all $\alpha > 0$, $\phi \in (0, 2\pi)$. Taking the supremum over ϕ, gives

$$\theta \geqslant \mu \frac{k^2\alpha}{2(1+\alpha)^k - 1} \quad \text{for all } \alpha > 0.$$

Thus we have (2.21) and, consequently, the bound on θ given by (2.18). □

Section 2.5 provides numerical values of the bound (2.18) for $\mu = \frac{1}{2}$ and $2 \leqslant k \leqslant 9$.

2.5. *Discussion*

Table 1 displays numerical values (rounded to 3 decimal places) of the lower stability bounds (2.11), (2.16), (2.18) on θ for the CS scheme with $\sigma = \frac{1}{2}$, the MCS scheme with $\sigma = \theta$ and $\mu = \frac{1}{2} - \theta$, and the HV scheme with $\mu = \frac{1}{2}$, respectively.

For each of these three bounds on θ we have shown that it is *necessary* for unconditional stability in the case of general diffusion equations (1.5) in arbitrary spatial dimension k, cf. Theorems 2.3, 2.6, 2.9. Subsequently, we proved the main result that, for any given dimension $k \geqslant 2$, the bound (2.18) for the HV scheme is also *sufficient* for unconditional stability, cf. Theorem 2.8. For the CS and MCS schemes we showed the sufficiency of their respective stability bounds if $k = 2$ or $k = 3$, cf. Theorems 2.2, 2.5.

All three stability bounds grow essentially linearly with k. The rate of growth for the HV scheme is approximately 0.116, for the MCS scheme 0.134, and for the CS scheme 0.184.

For the CS scheme one clearly observes in Table 1 that $\theta > \frac{1}{2}$ whenever $k \geqslant 4$, so that second order is lost. For the HV and MCS schemes under consideration we have order 2 for any value of θ.

Initial numerical experiments have been performed in two and three spatial dimensions with all the ADI schemes discussed in this paper, and in these experiments the classical order was always observed as the actual order of convergence. Next, the size of the parameter θ appeared to determine the size of the error constant – smaller values of θ led to smaller error constants. It is our aim of future research to investigate these issues in detail.

Appendix A. Proof of Theorem 2.8

For $k \geqslant 1$ define functions g_k and h_k by

$$g_k(y_1, \ldots, y_k) = \frac{(\sum_{j=1}^k y_j)^2}{2\prod_{j=1}^k (1 + y_j^2) - 1} \tag{A.1}$$

and

$$h_k(\alpha) = g_k\big(\sqrt{\alpha}, \ldots, \sqrt{\alpha}\big) = \frac{k^2\alpha}{2(1+\alpha)^k - 1} \quad (\alpha \geqslant 0). \tag{A.2}$$

It is easily verified that $g_1(y_1) < \frac{1}{2}$ for all y_1. Subsequently, there holds

$$\frac{1}{2} < \max_{0 \leqslant \alpha \leqslant 1} h_2(\alpha) < \max_{0 \leqslant \alpha \leqslant 1} h_3(\alpha) < \max_{0 \leqslant \alpha \leqslant 1} h_4(\alpha) < \cdots. \tag{A.3}$$

The inequalities (A.3) are illustrated in Fig. 1. The first inequality follows directly upon taking $\alpha = 1$; the remaining inequalities of (A.3) will be derived in Section A.3.

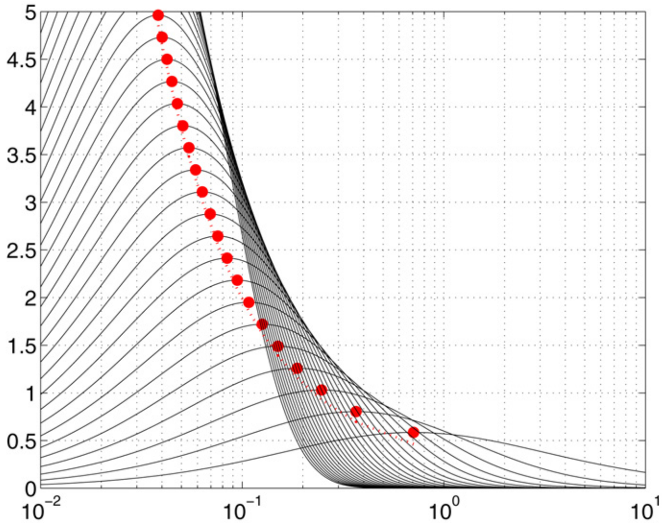

Fig. 1. Graphs of the functions h_k for increasing values of $k = 2, 3, \ldots, 30$ (bottom to top). The maximum of h_k is identified by a bullet. The dotted line forms an (under)estimate of the maximum of h_k based on asymptotic analysis.

A.1. *Proof of (2.21) ⇒ (2.20)*

We prove by induction on $\ell \geqslant 2$ that

$$g_\ell(y_1, \ldots, y_\ell) \leqslant \max_{0 \leqslant \alpha \leqslant 1} h_\ell(\alpha) \quad \text{whenever } y_1, \ldots, y_\ell \geqslant 0. \tag{A.4}$$

It is straightforward to show that (A.4) is true when $\ell = 2$. Let $k \geqslant 3$ be given and suppose (A.4) holds for $2 \leqslant \ell \leqslant k-1$.

(a) *The boundary of the domain.* For $1 \leqslant \ell \leqslant k-1$ we have

$$g_k(y_1, \ldots, y_\ell, 0, \ldots, 0) = g_\ell(y_1, \ldots, y_\ell).$$

If $l = 1$ then, using (A.3),

$$g_k(y_1, \ldots, y_\ell, 0, \ldots, 0) < \frac{1}{2} < \max_{0 \leqslant \alpha \leqslant 1} h_k(\alpha).$$

If $2 \leqslant \ell \leqslant k-1$ then, using (A.4) and (A.3),

$$g_k(y_1, \ldots, y_\ell, 0, \ldots, 0) \leqslant \max_{0 \leqslant \alpha \leqslant 1} h_\ell(\alpha) < \max_{0 \leqslant \alpha \leqslant 1} h_k(\alpha).$$

Next, for $0 \leqslant \ell \leqslant k-1$, we obtain

$$\begin{aligned}
\lim_{y_{\ell+1}, \ldots, y_k \to \infty} g_k(y_1, \ldots, y_k) &= \lim_{w_{\ell+1}, \ldots, w_k \downarrow 0} g_k\left(y_1, \ldots, y_\ell, \frac{1}{w_{\ell+1}}, \ldots, \frac{1}{w_k}\right) \\
&= \lim_{w_{\ell+1}, \ldots, w_k \downarrow 0} \frac{(q \sum_{j=1}^{\ell} y_j + s)^2}{2r \prod_{j=1}^{\ell} (1 + y_j^2) - q^2} \\
&= \begin{cases} \frac{1}{2 \prod_{j=1}^{k-1} (1 + y_j^2)} \leqslant \frac{1}{2} & \text{if } \ell = k-1, \\ 0 & \text{if } 0 \leqslant \ell < k-1 \end{cases} \\
&< \max_{0 \leqslant \alpha \leqslant 1} h_k(\alpha),
\end{aligned}$$

where

$$q = \prod_{j=\ell+1}^{k} w_j, \quad r = \prod_{j=\ell+1}^{k} (1 + w_j^2), \quad s = \sum_{j=\ell+1}^{k} \prod_{i=\ell+1, i \neq j}^{k} w_i,$$

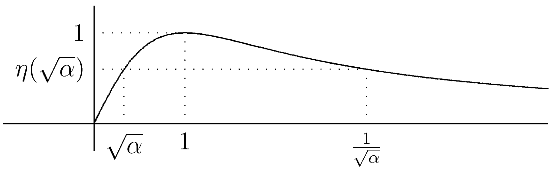

Fig. 2. The function η.

with the convention that empty sums are equal to zero and empty products equal to one. It thus follows, with a symmetry argument, that (A.4) is true for $\ell = k$ whenever $(y_1, \dots, y_k)$ is on the boundary of the domain $[0, \infty)^k$.

(b) *Critical points.* Interior critical points of g_k satisfy

$$\eta(y_i) := \frac{2y_i}{1+y_i^2} = \frac{1}{\sum_{j=1}^k y_j}\left(2 - \frac{1}{\prod_{j=1}^k (1+y_j^2)}\right), \quad i = 1, \dots, k.$$

Therefore, if $(y_1, \dots, y_k) \in (0, \infty)^k$ is a critical point of g_k, then

$$\eta(y_1) = \cdots = \eta(y_k) = \eta\bigl(\sqrt{\alpha}\bigr)$$

for some $0 < \alpha \leqslant 1$ (cf. Fig. 2), so that $y_i = \sqrt{\alpha}$ or $y_i = \frac{1}{\sqrt{\alpha}}$.

Let $0 \leqslant m \leqslant k$ be the number of values $y_i = \frac{1}{\sqrt{\alpha}}$. Then

$$g_k(y_1, \dots, y_k) = h_{k,m}(\alpha)$$

with

$$h_{k,m}(\alpha) := \frac{[(k-m)\sqrt{\alpha} + m/\sqrt{\alpha}]^2}{2(1+\alpha)^{k-m}(1+1/\alpha)^m - 1} = \frac{\alpha^{m-1}[(k-m)\alpha + m]^2}{2(1+\alpha)^k - \alpha^m} > 0.$$

It turns out that

$$\sup_{0<\alpha\leqslant 1} h_{k,0}(\alpha) \geqslant \sup_{0<\alpha\leqslant 1} h_{k,1}(\alpha) \geqslant \cdots \geqslant \sup_{0<\alpha\leqslant 1} h_{k,k}(\alpha). \tag{A.5}$$

The inequalities (A.5) are proved in Section A.4 and also illustrated in Fig. 3. Hence, if $(y_1, \dots, y_k) \in (0, \infty)^k$ is a critical point of g_k, then

$$g_k(y_1, \dots, y_k) \leqslant \sup_{0<\alpha\leqslant 1} h_{k,m}(\alpha) \leqslant \sup_{0<\alpha\leqslant 1} h_{k,0}(\alpha) = \max_{0\leqslant\alpha\leqslant 1} h_k(\alpha).$$

In view of (a) and (b) above, we obtain that (A.4) holds for $\ell = k$, and this concludes the induction proof.

Remark A.1. It is tempting, using the Cauchy–Schwarz inequality, to replace condition (2.20) by the stronger condition

$$\theta \geqslant \mu \frac{k\sum_{j=1}^k y_j^2}{2\prod_{j=1}^k (1+y_j^2) - 1} =: \mu \tilde{g}_k(y_1, \dots, y_k). \tag{A.6}$$

After all, the conditions (2.20) and (A.6) are identical for $y_1 = \cdots = y_k$. The analysis of interior critical points would also be much simpler under (A.6). However, $\tilde{g}_k(y_1, \dots, y_\ell, 0, \dots, 0) \neq \tilde{g}_\ell(y_1, \dots, y_\ell)$. In fact, $\tilde{g}_k(y_1, 0, \dots, 0) = ky_1^2/(1+2y_1^2) \to k/2$ as $y_1 \to \infty$ (a situation where the Cauchy–Schwarz inequality used above is far from tight), so that (A.6) would at best lead to $a_k = \frac{1}{2}$ in Theorem 2.8, whereas for a_k given by (2.19) we always have $a_k \in (0.20, 0.30)$.

A.2. Proof of (2.21) ⇔ (2.18)

Let $k \geqslant 2$. The function h_k satisfies

$$h_k'(\alpha) = k^2 \frac{2[1-(k-1)\alpha](1+\alpha)^{k-1} - 1}{[2(1+\alpha)^k - 1]^2} \quad (\alpha \geqslant 0).$$

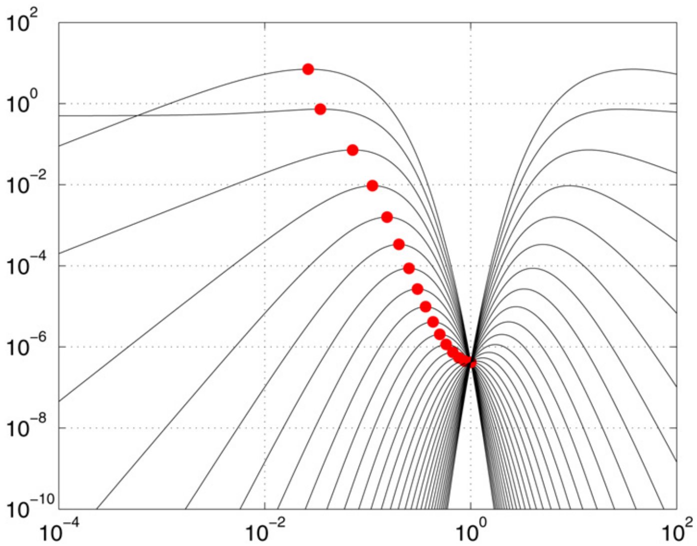

Fig. 3. Logarithmic plot of the functions $h_{k,m}$ for $0 \leqslant m \leqslant k = 30$ ("top" to bottom with increasing m on the left of $\alpha = 10^0$, with decreasing m on the right). The bullets identify the maxima of $h_{k,m}$ on (0, 1].

Clearly, $h_k'(\frac{1}{2(k-1)}) > 0$ and $h_k'(\frac{1}{k-1}) < 0$. Further, it is easily verified that the numerator of h_k' is a decreasing function of α. Hence, h_k' has a unique zero α_k with $\frac{1}{2} < (k-1)\alpha_k < 1$, corresponding to an absolute maximum of h_k. Define $a_k = 1-(k-1)\alpha_k$. Then $a = a_k$ satisfies (2.19) and we have

$$h_k(\alpha_k) = \frac{k^2\alpha_k}{2(1+\alpha_k)^k - 1} = ka_k.$$

A.3. *Proof of (A.3)*

Let $k \geqslant 2$. As seen above, the function h_k attains its absolute maximum on the interval $(0, \frac{1}{k-1})$. We will prove below that h_{k+1} is strictly greater than h_k in this interval. Define

$$\delta_k(\alpha) = 2(k+1)^2(1+\alpha)^k - 2k^2(1+\alpha)^{k+1} - (k+1)^2 + k^2.$$

Then for $\alpha > 0$ we have

$$h_{k+1}(\alpha) > h_k(\alpha) \iff \delta_k(\alpha) > 0.$$

We consider δ_k. First,

$$\begin{aligned}\delta_k'(\alpha) &= 2k(k+1)^2(1+\alpha)^{k-1} - 2k^2(k+1)(1+\alpha)^k \\ &= 2k(k+1)(1+\alpha)^{k-1}(1-k\alpha).\end{aligned}$$

Thus δ_k increases on $[0, \frac{1}{k}]$ and decreases on $[\frac{1}{k}, \infty)$. Further, $\delta_k(0) = 2k+1 > 0$. In the following we will show that also $\delta_k(\frac{1}{k-1}) > 0$, which yields that $\delta_k(\alpha) > 0$ whenever $0 < \alpha < \frac{1}{k-1}$ and concludes the proof. Accordingly, for $k\alpha \leqslant 2$ there holds

$$\begin{aligned}\delta_k(\alpha) &= 2(1+\alpha)^k\big[(k+1)^2 - k^2(1+\alpha)\big] - (k+1)^2 + k^2 \\ &= 2(1+\alpha)^k\big[(2-k\alpha)k+1\big] - (2k+1) \\ &\geqslant 2(1+k\alpha)\big[(2-k\alpha)k+1\big] - (2k+1) \\ &= 2\big[(1+k\alpha)(2-k\alpha)-1\big]k + 1 + 2k\alpha \\ &= 2\big[1+k\alpha-(k\alpha)^2\big]k + 1 + 2k\alpha.\end{aligned}$$

If $k \geqslant 3$ and $\alpha = \frac{1}{k-1}$, then $1 + k\alpha - (k\alpha)^2 > 0$ and we obtain $\delta_k(\alpha) \geqslant 1 > 0$. By direct substitution it is readily verified that $\delta_k(\frac{1}{k-1}) > 0$ also for $k = 2$.

A.4. Proof of (A.5)

Let $k \geqslant 2$ and $0 < \alpha \leqslant 1$ and consider m as a real variable. For the logarithmic derivative of $h_{k,m}(\alpha)$ with respect to $m \in [1, k]$ there holds

$$\begin{aligned}\frac{\partial}{\partial m}\ln\bigl[h_{k,m}(\alpha)\bigr] &= \ln\alpha + \frac{2(1-\alpha)}{(k-m)\alpha+m} + \frac{\alpha^m \ln\alpha}{2(1+\alpha)^k - \alpha^m}\\ &= \frac{2\ln\alpha}{2-\alpha^m(1+\alpha)^{-k}} + \frac{2(1-\alpha)}{(k-m)\alpha+m}\\ &\leqslant \ln\alpha + \frac{2(1-\alpha)}{k}\frac{1}{1-(k-m)(1-\alpha)/k}\\ &= \sum_{n\geqslant 1} c_n(k,m)\left(\frac{1-\alpha}{k}\right)^n\end{aligned}$$

where

$$c_n(k,m) = -\frac{k^n}{n} + 2(k-m)^{n-1} \quad (n \geqslant 1).$$

We will show that $c_n(k,m) \leqslant 0$ for all $k \geqslant 2$, $1 \leqslant m \leqslant k$, $n \geqslant 1$. Clearly,

$$c_n(k,m) \leqslant c_n(k,1) = -\frac{k^n}{n} + 2(k-1)^{n-1} =: d_n(k)$$

whenever $1 \leqslant m \leqslant k$. Consider k as a real variable. First, we have $d_1(k) = 2 - k \leqslant 0$ $(k \geqslant 2)$. Next suppose that for given $n \geqslant 2$ there holds $d_{n-1}(k) \leqslant 0$ $(k \geqslant 2)$. Since $d_n(2) = -\frac{1}{n}2^n + 2 \leqslant 0$ and $d_n'(k) = (n-1)d_{n-1}(k) \leqslant 0$ we obtain $d_n(k) \leqslant 0$ $(k \geqslant 2)$. It follows that all coefficients $d_n(k)$, and therefore $c_n(k,m)$, are non-positive. This implies that $h_{k,1}(\alpha) \geqslant h_{k,2}(\alpha) \geqslant \cdots \geqslant h_{k,k}(\alpha)$ for every $0 < \alpha \leqslant 1$, which proves the last $k-1$ inequalities in (A.5).

The first inequality in (A.5) needs special treatment because $h_{k,0}$ and $h_{k,1}$ intersect for $0 < \alpha < 1$, as evidenced by Fig. 3. Since

$$\frac{h_{k,1}'(\alpha)}{h_{k,1}(\alpha)} = \frac{[1-(k-1)\alpha][1+2(k-2)(1+\alpha)^{k-1}]}{[(k-1)\alpha+1][2(1+\alpha)^k - \alpha]} \quad (\alpha \geqslant 0)$$

the maximum of $h_{k,1}$ occurs at $\alpha = \frac{1}{k-1}$, where the inequality $h_{k,0}(\alpha) \geqslant h_{k,1}(\alpha)$ is easily seen to hold by comparing respective numerators and denominators. This completes the proof of (A.5).

References

[1] P.L.T. Brian, A finite-difference method of high-order accuracy for the solution of three-dimensional transient heat conduction problems, AIChE J. 7 (1961) 367–370.

[2] I.J.D. Craig, A.D. Sneyd, An alternating-direction implicit scheme for parabolic equations with mixed derivatives, Comp. Math. Appl. 16 (1988) 341–350.

[3] J. Douglas, Alternating direction methods for three space variables, Numer. Math. 4 (1962) 41–63.

[4] J. Douglas, H.H. Rachford, On the numerical solution of heat conduction problems in two and three space variables, Trans. Amer. Math. Soc. 82 (1956) 421–439.

[5] D.J. Duffy, Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach, Wiley, 2006.

[6] K.J. in 't Hout, B.D. Welfert, Stability of ADI schemes applied to convection–diffusion equations with mixed derivative terms, Appl. Numer. Math. 57 (2007) 19–35.

[7] W. Hundsdorfer, Accuracy and stability of splitting with stabilizing corrections, Appl. Numer. Math. 42 (2002) 213–233.

[8] W. Hundsdorfer, J.G. Verwer, Numerical Solution of Time-Dependent Advection–Diffusion–Reaction Equations, Springer Series in Computational Mathematics, vol. 33, Springer-Verlag, Berlin, 2003.

[9] D.W. Peaceman, Fundamentals of Numerical Reservoir Simulation, Developments in Petroleum Science, vol. 6, Elsevier, Amsterdam, 1977.

[10] D. Tavella, C. Randall, Pricing Financial Instruments: The Finite Difference Method, Wiley, New York, 2000.