NUMERICAL SOLUTION OF PROBLEMS WITH DIFFERENT TIME SCALES II

By

Heinz-Otto Kreiss

IMA Preprint Series # 1294

February 1995

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

514 Vincent Hall

206 Church Street S.E.

Minneapolis, Minnesota 55455

Numerical solution of problems with different time scales II.

by

Heinz-Otto Kreiss

1. Introduction

Consider the initial value problem for a system of differential equations

$$
\begin{aligned}
&u_t = \varepsilon^{-1} Au + Bu + F, \quad t \geq 0, \\
&u(0) = u_0.
\end{aligned} \tag{1.1}
$$

Here $u = (u_1, \ldots, u_n)^T$, $F = F(t,\varepsilon) = (F_1(t,\varepsilon), \ldots, F_n(t,\varepsilon))^T \in C^\infty(t,\varepsilon)$ are vector functions with n components and $A = A(t,\varepsilon) \in C^\infty(t,\varepsilon)$, $B = B(t,\varepsilon) \in C^\infty(t,\varepsilon)$ are $n \times n$ matrices which depend on t and a small parameter ε for $0 < \varepsilon \leq \varepsilon_0 << 1$. We assume also that A, B, F vary slowly as functions of t. To make this precise we define

Definition 1.1. *A function $f(t,\varepsilon)$ is said to vary slowly of order p if f and its derivatives $d^j f/dt^j$, $j \leq p$ can be estimated independently of ε. If such an estimate holds for any p, then we say that the function varies slowly.*

We are not interested in problems with rapidly growing solutions where the large matrix is degenerate like

$$
u_t = \left(\frac{1}{\varepsilon} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{pmatrix} \right) u.
$$

Therefore, we make

Assumption 1.1. *There is a slowly varying transformation $S(t,\varepsilon)$ such that*

$$
S^{-1}(t,\varepsilon) A(t,\varepsilon) S(t,\varepsilon) = \begin{pmatrix} A_{11}(t,\varepsilon) & 0 \\ 0 & \varepsilon A_{22}(t,\varepsilon) \end{pmatrix}, \tag{1.2}
$$

where A_{11}, A_{22} are slowly varying and

$$
(1.3) \qquad A_{11} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}, \quad |\lambda_j| \geq 1, \quad j = 1, 2, \ldots, m.
$$

In this paper we are only interested in the highly oscillatory case and, therefore, we make

Assumption 1.2. *The eigenvalues of A_{11} are purely imaginary, i.e.,*

$$
\text{Re } \lambda_j = 0, \quad j = 1, 2, \ldots, m.
$$

The real stiff case is treated in [2].

An important special case arises when $A_{22} \equiv 0$. To achieve this goal one has to be careful how one splits a given matrix into its large part $\frac{1}{\varepsilon}A$ and its small part B. An example is

$$
u_t = \left(i \begin{pmatrix} -\frac{1}{\varepsilon} & \frac{1}{\varepsilon} \\ \frac{1}{\varepsilon} & -(\frac{1}{\varepsilon} + 1) \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \right) u =: (\frac{1}{\varepsilon}A + B)u.
$$

The elements of $\frac{1}{\varepsilon}A$ are all large but the rows are almost linearly dependent. The eigenvalues of $\frac{1}{\varepsilon}A$ are

$$
\lambda_1 = -\frac{2i}{\varepsilon} + \mathcal{O}(1), \quad \lambda_2 = -\frac{i}{2} + \mathcal{O}(\varepsilon)
$$

and, therefore, $A_{22} \neq 0$. The remedy is to write the system in the form

$$
u_t = \left(i \begin{pmatrix} -\frac{1}{\varepsilon} & \frac{1}{\varepsilon} \\ \frac{1}{\varepsilon} & -\frac{1}{\varepsilon} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} - i \end{pmatrix} \right) u.
$$

There is one other situation we have to be aware of. Consider

$$u_t = \left(\begin{pmatrix} 0 & \frac{1}{\varepsilon^2} \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{pmatrix} \right) u \tag{1.4}$$

which could also be written as

$$u_t = \left(\begin{pmatrix} 0 & \frac{1}{\varepsilon^2} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ -1 & b_{22} \end{pmatrix} \right) u.$$

Both versions violate our assumption. Instead, we have to "rescale" the system by introducing new dependent variables

$$\varepsilon u_1 = u_1', \quad u_2 = u_2'.$$

Now (1.4) becomes

$$u_t' = \left(\frac{1}{\varepsilon} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} b_{11} & \varepsilon b_{12} \\ 0 & b_{22} \end{pmatrix} \right) u',$$

and our assumptions are satisfied. This behavior is very common when discretizing hyperbolic problems with different time scales.

The plan of our paper is as follows. In Section 2 we collect analytic results for the solution of (1.1). In Section 3 we shall solve our problem by asymptotic expansions. In this paper we are not interested in highly oscillatory solutions. Therefore, in Section 4 we shall discuss methods to "initialize the data" such that the resulting solution varies only on the slow time scale.

In Section 5 we will discuss the backward differentiation methods (BDF)

$$\left(I - \frac{2}{3} k \left(\frac{1}{\varepsilon} A_{n+1} + B_{n+1} \right) \right) y_{n+1} = \sum_{j=0}^{p-1} \alpha_j y_{n-j} + k F_{n+1}$$

and, in Section 6, apply them to hyperbolic partial differential equations.

In many applications it is much easier to invert $I - \frac{2}{3}\frac{k}{\varepsilon}A$ than $I - \frac{2}{3}k(\frac{1}{\varepsilon}A + B)$. Therefore, in the last two sections, we consider split methods

$$(1.5) \qquad (I - \frac{2}{3}\frac{k}{\varepsilon}A_{n+1})y_{n+1} = \sum_{j=0}^{p-1} \alpha_j y_{n-j} + \sum_{j=0}^{r-1} \beta_j (B_{n-j}y_{n-j} + F_{n-j}).$$

Our results indicate that the best methods are

1) The classic combination of Leap-frog and Crank-Nicolson.

2) A combination of the second order BDF method with an explicit second or third order method of type (1.5).

In the literature one can find other splitting techniques like "Strang" and additative splitting. In our opinion, they cannot compete with split multistep methods. (For details, see [1]).

2. Analytic results

We use Assumption 1.1 to introduce a new variable $u_1 = S^{-1}u$ into (1.1) and obtain

$$u_{1t} = \left(\begin{pmatrix} \frac{1}{\varepsilon}A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} + B^{(1)} \right) u_1 + F^{(1)}, \tag{2.1}$$

where

$$B^{(1)} = S^{-1}BS - S^{-1}S_t, \quad F^{(1)} = S^{-1}F.$$

There is a slowly varying transformation $I + \varepsilon S_1$ such that

$$(I + \varepsilon S_1)^{-1} \left(\begin{pmatrix} \frac{1}{2}A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} + B^{(1)} \right) (I + \varepsilon S_1)$$
$$= \begin{pmatrix} \frac{1}{\varepsilon}(A_{11} + \varepsilon C_{11}^{(1)}) & 0 \\ 0 & A_{22} + \varepsilon C_{22}^{(1)} \end{pmatrix} =: D.$$

Here $C_{11}^{(1)}$ is diagonal. Therefore, $u_2 = (I + \varepsilon S_1)^{-1}u_1$ satisfies

$$u_{2t} = (D + \varepsilon B^{(2)})u_2 + F^{(2)}.$$

We can repeat the process and after $p + 1$ transformations we obtain a system of the form

$$\begin{pmatrix} u^I \\ u^{II} \end{pmatrix}_t = \begin{pmatrix} \frac{1}{\varepsilon}(A_{11} + \varepsilon C_{11}) & \varepsilon^p C_{12} \\ \varepsilon^p C_{21} & B_{22} + \varepsilon C_{22} \end{pmatrix} \begin{pmatrix} u^I \\ u^{II} \end{pmatrix} + F_2.$$

Thus, except for terms of order ε^p, we obtain the decoupled system

$$u_t^I = \frac{1}{\varepsilon}(A_{11} + \varepsilon C_{11})u^I + F^I, \quad C_{11} \text{ diagonal,} \tag{2.2a}$$

$$u_t^{II} = (B_{22} + \varepsilon C_{22})u^{II} + F^{II}. \tag{2.2b}$$

We shall now derive an asymptotic expansion for the solutions of (2.2a). We want to show that it consists of a slowly varying and a rapidly varying part.

If we neglect u_t^I, then $u^I = -\varepsilon(A_{11} + \varepsilon C_{11})^{-1} F^I$. Therefore, we make the substitution

$$u^I = \varepsilon\varphi_1 + u_1^I, \quad \varphi_1 = -(A_{11} + \varepsilon C_{11})^{-1} F$$

and obtain

$$u_{1t}^I = \frac{1}{\varepsilon}(A_{11} + \varepsilon C_{11})u^I + \varepsilon F_1^I, \quad F_1^I = \Big((A_{11} + \varepsilon C_{11})^{-1} F\Big)_t.$$

Thus, we have reduced the forcing to order $\mathcal{O}(\varepsilon)$. After p steps we obtain

$$u^I = \sum_{j=1}^{p} \varepsilon^j \varphi_j + u_p^I, \tag{2.3}$$

where

$$u_{pt}^I = \frac{1}{\varepsilon}(A_{11} + \varepsilon C_{11})u_p^I + \varepsilon^p F_p^I,$$
$$u_p^I(0) = u^I(0) - \sum_{j=1}^{p} \varepsilon^j \varphi_j(0).$$

If the initial data are such that $u_p^I(0) \equiv \mathcal{O}(\varepsilon^p)$, then $u_p^I(t) = \mathcal{O}(\varepsilon^p)$ and has p derivatives bounded independently of ε. Otherwise, it is dominated by the rapidly varying part

$$u_p^I(t) = e^{\frac{1}{\varepsilon}\int_0^t (A_{11} + \varepsilon C_{11})d\xi} u_p^I(0). \tag{2.4}$$

Since the solutions of (2.2b) are slowly varying we have sketched a proof of (for more details, see[4],[5] and [6])

Theorem 2.1. *The solution of (1.1) can be written as*

$$u = u_S + u_F.$$

Here u_S is slowly and u_F is rapidly varying. Any initial data $u(0)$ can be split into

$$u(0) = u_S(0) + u_F(0).$$

$u_S(0)$ generates the slowly varying and $u_F(0)$ the rapidly varying part. In particular, $u_S(t)$ is determined by (2.2b) and (2.3) with $u_p^I \equiv 0$. Also, u_S is a smooth function of ε.

3. Numerical solution by asymptotic expansion

To use asymptotic expansions to solve (1.1) we need to calculate the transformation (1.3) efficiently. Techniques to do so are described in [5]. We assume that we have executed this transformation and that the equations are given in the form

$$u_t^I = \varepsilon^{-1}(A_{11} + \varepsilon B_{11})u^I + B_{12}u^{II} + F^I, \tag{3.1a}$$

$$u_t^{II} = (A_{22} + B_{22})u^{II} + B_{21}u^I + F^{II}. \tag{3.1b}$$

For slowly varying solutions, we can, to first approximation, neglect u_t^I. Therefore, we solve

$$\begin{aligned} &(A_{11} + \varepsilon B_{11})\psi_0^I + \varepsilon B_{12}\psi_0^{II} + \varepsilon F^I = 0, \\ &\psi_{0t}^{II} = (A_{22} + B_{22})\psi_0^{II} + B_{21}\psi_0^I + F^{II}, \\ &\psi_0^{II}(0) = u^{II}(0). \end{aligned} \tag{3.2}$$

The remainder

$$u_1 = u - \psi_0$$

is a solution of

$$\begin{aligned} &u_{1t}^I = \varepsilon^{-1}(A_{11} + \varepsilon B_{11})u_1^I + B_{12}u_1^{II} - \psi_{0t}^I, \\ &u_{1t}^{II} = (A_{22} + \varepsilon B_{22})u_1^{II} + B_{21}u_1^I, \\ &u_1^{II}(0) = 0. \end{aligned} \tag{3.3}$$

Now we solve

$$\begin{aligned} &(A_{11} + \varepsilon B_{11})\psi_1^I + \varepsilon B_{12}\psi_1^{II} - \varepsilon^{-1}\psi_{0t}^I = 0, \\ &\psi_{1t}^{II} = (A_{22} + B_{22})\psi_1^{II} + B_{21}\psi_1^I, \\ &\psi_1^{II}(0) = 0. \end{aligned} \tag{3.4}$$

(Observe that $\psi_{0t}^I = \mathcal{O}(\varepsilon)$ and, therefore, $\psi_1^I = \mathcal{O}(1)$.) Then

$$u_2 = u_1 - \varepsilon^2\psi_1$$

is a solution of

$$
(3.5)\qquad
\begin{aligned}
u^I_{2t} &= \varepsilon^{-1}(A_{11}+\varepsilon B_{11})u^I_2 + B_{12}u^{II}_2 - \varepsilon^2\psi^I_{1t},\\
u^{II}_{2t} &= (A_{22}+\varepsilon B_{22})u^{II}_2 + B_{21}u^I_2,\\
u^{II}_2(0) &= 0.
\end{aligned}
$$

The reduction process can be continued. After p steps we obtain

$$
u = \psi^{(p)} + w, \quad \psi^{(p)} = \psi_0 + \varepsilon \sum_{j=1}^{p-1} \varepsilon^j \psi_j,
$$

where, except for terms of order $\mathcal{O}(\varepsilon^{p+1})$, w solves

$$
(3.6)\qquad
\begin{aligned}
&\begin{pmatrix} w^I \\ w^{II} \end{pmatrix}_t = \begin{pmatrix} \frac{1}{\varepsilon}(A_{11}+\varepsilon B_{11}) & B_{12} \\ B_{21} & A_{22}+B_{22} \end{pmatrix} \begin{pmatrix} w^I \\ w^{II} \end{pmatrix} =: \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} w^I \\ w^{II} \end{pmatrix}\\
&w^I(0) = u^I(0) - (\psi^{(p)}(0))^I, \quad w^{II}(0) = 0.
\end{aligned}
$$

If $w^I(0) = 0$, then $w \equiv 0$ and we have solved our problem, which, in this case, has no fast part. If $w^I(0) \neq 0$, then we have to separate the scales further. We make a variable substitution of type

$$
w = \begin{pmatrix} I & S \\ 0 & I \end{pmatrix} \tilde{w}
$$

and obtain

$$
\begin{aligned}
\tilde{w}_t + \begin{pmatrix} 0 & S_t \\ 0 & 0 \end{pmatrix} \tilde{w} &= \begin{pmatrix} I & -S \\ 0 & I \end{pmatrix} \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} I & S \\ 0 & I \end{pmatrix} \tilde{w}\\
&= \begin{pmatrix} D_{11} - SD_{21} & D_{12} - SD_{22} \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} I & S \\ 0 & I \end{pmatrix} \tilde{w}\\
&= \begin{pmatrix} D_{11} - SD_{21} & D_{11}S - SD_{22} + D_{12} - SD_{21}S \\ D_{21} & D_{21}S + D_{22} \end{pmatrix} \tilde{w}.
\end{aligned}
$$

Thus,

$$
(3.7)\qquad
\begin{aligned}
\tilde{w}_t &= \left(\begin{pmatrix} D_{11} - SD_{21} & 0 \\ D_{21} & D_{21}S + D_{22} \end{pmatrix} + \mathcal{O}(\varepsilon^{p+1}) \right) \tilde{w}.\\
\tilde{w}^I(0) &= w^I(0), \quad \tilde{w}^{II}(0) = 0,
\end{aligned}
$$

if we choose S such that

$$
\begin{aligned}
S_t &= D_{11}S - SD_{22} + D_{12} - SD_{21}S + \mathcal{O}(\varepsilon^{p+1}) \\
&= \frac{1}{\varepsilon}(A_{11} + \varepsilon B_{11})S - S(A_{22} + B_{22}) + B_{12} - SB_{21}S + \mathcal{O}(\varepsilon^{p+1}).
\end{aligned} \tag{3.8}
$$

(3.8) is satisfied if we use $p+1$ steps of the iteration

$$
\varepsilon S_t^{(j)} = A_{11}S^{(j+1)} + \varepsilon B_{12} + \varepsilon B_{11}S^{(j)} - \varepsilon S^{(j)}(A_{22} + B_{22}) - \varepsilon S^{(j)}B_{21}S^{(j)}
$$

$$
S^{(1)} = -\varepsilon A_{11}^{-1}B_{12}.
$$

Now we neglect the term $\mathcal{O}(\varepsilon^{p+1})$ in (3.7) and make the substitution

$$
\tilde{w} = \begin{pmatrix} I & 0 \\ T & I \end{pmatrix} \tilde{\tilde{w}}.
$$

In the same way as above, neglecting terms of order $\mathcal{O}(\varepsilon^{p+1})$, we obtain

$$
\begin{aligned}
\tilde{\tilde{w}}_t &= \begin{pmatrix} D_{11} - SD_{21} & 0 \\ 0 & D_{21}S + D_{22} \end{pmatrix} \tilde{\tilde{w}} \\
&= \begin{pmatrix} \frac{1}{\varepsilon}(A_{11} + \varepsilon B_{11}) - SB_{21} & 0 \\ 0 & A_{22} + B_{22} + B_{21}S \end{pmatrix} \tilde{\tilde{w}} \\
\tilde{\tilde{w}}^I(0) &= w^I(0), \quad \tilde{\tilde{w}}^{II}(0) = Tw^I(0),
\end{aligned} \tag{3.9}
$$

if we choose T such that

$$
\begin{aligned}
T_t &= -T(D_{11} - SD_{21}) + D_{21} + (D_{21}S + D_{22})T + \mathcal{O}(\varepsilon^{p+1}) \\
&= -T\big(\frac{1}{\varepsilon}(A_{11} + \varepsilon B_{11}) - SB_{21}\big) + B_{21} + (B_{21}S + A_{22} + B_{22})T + \mathcal{O}(\varepsilon^{p+1}).
\end{aligned} \tag{3.10}
$$

Again, we can determine T by iteration starting with

$$
T^{(1)} = \varepsilon B_{21}A_{11}^{-1}. \tag{3.11}
$$

As before, the system for $\tilde{\tilde{w}}^I$ can be solved by analytic techniques. To determine $\tilde{\tilde{w}}^{II}$ we can use a standard multistep or Runge-Kutta type method.

Except for terms of order $\mathcal{O}(\varepsilon^{p+1})$ we have separated the scales. The slow and the fast parts of the solution of (3.9) are obtained by choosing $\tilde{\tilde{w}}^I(0) = 0$ and $\tilde{\tilde{w}}^{II}(0) = 0$, respectively. Since

$$w = \begin{pmatrix} I & S \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ T & I \end{pmatrix} \tilde{\tilde{w}} = \begin{pmatrix} I+ST & S \\ T & I \end{pmatrix} \tilde{\tilde{w}},$$

we have, therefore,

$$(3.12) \quad \begin{aligned} u_S^I &= (\psi^{(p)})^I + w_S^I + \mathcal{O}(\varepsilon^{p+1}) = (\psi^{(p)})^I + S\tilde{\tilde{w}}^{II} + \mathcal{O}(\varepsilon^{p+1}), \\ u_S^{II} &= (\psi^{(p)})^{II} + w_S^{II} + \mathcal{O}(\varepsilon^{p+1}) = (\psi^{(p)})^{II} + \tilde{\tilde{w}}^{II} + \mathcal{O}(\varepsilon^{p+1}), \end{aligned}$$

$$(3.13) \quad \begin{aligned} u_F^I &= w_F^I + \mathcal{O}(\varepsilon^{p+1}) = (I+ST)\tilde{\tilde{w}}^I + \mathcal{O}(\varepsilon^{p+1}), \\ u_F^{II} &= w_F^{II} + \mathcal{O}(\varepsilon^{p+1}) = T\tilde{\tilde{w}}^I + \mathcal{O}(\varepsilon^{p+1}), \end{aligned}$$

Observe that $S = \mathcal{O}(\varepsilon)$, $T = \mathcal{O}(\varepsilon)$, $\tilde{\tilde{w}}^{II} = \mathcal{O}(\varepsilon)$ and, therefore,

$$u_S^I = (\psi^{(0)})^I + \mathcal{O}(\varepsilon^2), \quad u_S^{II} = (\psi^{(0)})^{II} + \varepsilon\varphi^{II} + \mathcal{O}(\varepsilon^2)$$

where

$$\varphi_t^{II} = (A_{22} + B_{22})\varphi^{II}, \quad \varphi^{II}(0) = B_{21}A_{11}^{-1}u^I(0).$$

4. Initialization

If the solution of our problem varies also on the fast scale, then we must either separate the scales as explained in the previous section or use a standard explicit multistep or Runge-Kutta method directly with a timestep $k << \varepsilon$.

In many applications one is not interested in the fast scale. Often, the energy contained in the fast scale solution is small and might also be introduced through observational errors. Therefore, one initializes the data, i.e., one changes the the initial data such that the resulting solution varies only on the slow scale. There are different method to accomplish this.

1) *Separation of scales.* We use the technique shown in the previous section to obtain initialized data $u_J^I(0)$, $u_J^{II}(0)$. The easiest way is to use

$$\begin{aligned} u_J^I(0) &= \psi_0^I(0) \approx -\varepsilon A_{11}^{-1}(B_{12}u^{II}(0) + F^I(0)), \\ u_J^{II}(0) &= \psi_0^{II}(0) = u^{II}(0). \end{aligned} \tag{4.1}$$

If the fast scale is only introduced through observational errors, we commit an error of order $\mathcal{O}(\varepsilon^2)$. If the fast scale is "real", then we should use a more accurate initialization. Neglecting terms of order $\mathcal{O}(\varepsilon^2)$, we obtain

$$\begin{aligned} u_J^I(0) &= \psi_0^I(0), \\ u_J^{II}(0) &= \psi_0^{II}(0) + \tilde{\tilde{w}}^{II}(0) = u^{II}(0) + \varepsilon B_{21}A_{11}^{-1}u^I(0). \end{aligned} \tag{4.2}$$

The term $\varepsilon B_{21}A_{11}^{-1}u^I(0)$ represents the effect of the fast scale on the slow scale solution.

2) *The bounded derivative principle.* We determine the initial data such that p time derivatives of the solution are bounded independently of ε at $t = 0$. Typically $p = 2$, and then it is equivalent with (4.1). For details, see [5].

3) *Richardson extrapolation.* Let $k > 0$ denote a timestep which defines gridpoints $t_\nu = \nu k,\ \nu = 0, 1, 2, \ldots$ and gridfunctions $y_\nu = y(\nu k)$. We approximate (1.1) by

$$(4.3)\qquad \begin{aligned} &(I - \frac{k}{\varepsilon}A)y(t_{n+1}) = (I + k\tilde{B})y(t_n) + kF(t_n),\\ &y(0) = u(0). \end{aligned}$$

Here $A = A(t_{n+1}, \varepsilon),\ \tilde{B} = B(t_n, \varepsilon) + kB^{(1)}(t_n, \varepsilon, k)$ are smooth functions of all variables.

Remark: In actual applications we use $\tilde{B} = B$. We have added $B^{(1)}$ to simplify the discussion below.

Our aim is to prove

Theorem 4.1. *Let $u_S = u_S(t, \varepsilon)$ be the smooth part of the solution of (1.1). For any integer $p > 0$, the solution $y = y(t_n, k, \varepsilon)$ of (4.3) can be expanded into a series*

$$(4.4)\qquad y(t_n, k, \varepsilon) = u_S(t_n, k, \varepsilon) + kE + \mathcal{O}(k^{p+1} + \varepsilon^{p+1} + (\varepsilon/k)^n).$$

Here kE stands for a Taylor expansion

$$(4.5)\qquad kE = ke_1(t_n, \varepsilon) + \cdots + k^p e_p(t_n, \varepsilon),$$

where $e_j(t, \varepsilon)$ are smooth functions of all variables. Therefore, for all n with $(\varepsilon/k)^n << k^{p+1}$, we can use Richardson extrapolation to calculate u_S accurately to order $\mathcal{O}(k^p)$.

The proof consists of a number of steps.

a) We use the transformation (1.2) to make a change of variables $y_{n+1} = S_{n+1}\tilde{y}_{n+1}$. Then we obtain an approximation of the same type for (2.1). Therefore, we can assume that A already has the blockdiagonal form (1.2).

b) Corresponding to the previous section, we can construct a particular slowly varying solution such that we can neglect the forcing F in (4.3). We solve

$$A_{11}\varphi_0^I(t_{n+1} + \varepsilon B_{11}\varphi_0^I(t_n) + \varepsilon \tilde{B}_{12}\varphi_0^{II}(t_n) + \varepsilon F^I(t_n) = 0,$$

$$(I - kA_{22})\varphi_0^{II}(t_{n+1}) = (I + k\tilde{B}_{22})\varphi_0^{II}(t_n) + k\tilde{B}_{21}\varphi_0^I(t_n) + F^{II}(t_n), \tag{4.6}$$

$$\varphi_0^{II}(0) = u^{II}(0).$$

(4.6) is a regular approximation of (3.2). Therefore, standard techniques show that there is a solution of the form

$$\varphi_0 = \psi_0 + kE_1 + \mathcal{O}(k^{p+1}), \tag{4.7}$$

where kE_1 represents a Taylor expansion of type (4.5). The remainder

$$y_1 = y - \varphi_0$$

is a solution of

$$(I - \frac{k}{\varepsilon}A_{11})y_1^I(t_{n+1}) = (I + k\tilde{B}_{11})y_1^I(t_n) + k\tilde{B}_{12}y_1^{II}(t_n) - (\varphi_0^I(t_{n+1} - \varphi_0^I(t_n)),$$

$$(I - kA_{22})y_1^{II}(t_{n+1}) = (I + k\tilde{B}_{22})y_1^{II}(t_n) + k\tilde{B}_{21}y_1^I(t_n), \tag{4.8}$$

$$y_1^{II}(0) = 0,$$

which is an approximation of (3.3). Observing that

$$\varphi_0^I(t_{n+1}) - \varphi_0^I(t_n) = \psi_{0t} + kE_1 + \mathcal{O}(k^{p+1}),$$

we repeat the construction and obtain

$$\varphi_1 = \psi_1 + kE_2 + \mathcal{O}(k^{p+1}).$$

This process can be continued and, therefore, we have proved

Lemma 4.1. *Corresponding to the procedure in the previous section, we can construct a solution*

$$\varphi^{(p)} = \psi^{(p)} + kE + \mathcal{O}(k^{p+1})$$

such that, neglecting terms of order $\mathcal{O}(\varepsilon^{p+1} + k^{p+1})$,

$$v = y - \varphi^{(p)}$$

solves the homogeneous system (4.8) which we write as

$$\begin{aligned} v(t_{n+1}) &= \begin{pmatrix} (I - \frac{k}{\varepsilon}A_{11})^{-1}(I + k\tilde{B}_{11}) & k(I - \frac{k}{\varepsilon}A_{11})^{-1}\tilde{B}_{12} \\ k(I - kA_{22})^{-1}\tilde{B}_{21} & (I - kA_{22})^{-1}(I + k\tilde{B}_{22}) \end{pmatrix} v(t_n) \\ &=: \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} v(t_n). \end{aligned}$$

We now introduce a transformation

$$v(t_n) = \begin{pmatrix} I & \tilde{S}_n \\ 0 & I \end{pmatrix} \tilde{v}(t_n)$$

and obtain

$$\begin{aligned} \tilde{v}(t_{n+1}) &= \begin{pmatrix} I & -\tilde{S}_{n+1} \\ 0 & I \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} I & \tilde{S}_n \\ 0 & I \end{pmatrix} \tilde{v}(t_n) \\ &= \begin{pmatrix} C_{11} - \tilde{S}_{n+1}C_{21} & C_{12} - \tilde{S}_{n+1}C_{22} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} I & \tilde{S}_n \\ 0 & I \end{pmatrix} \tilde{v}(t_n) \\ &= \begin{pmatrix} C_{11} - \tilde{S}_{n+1}C_{21} & \tilde{C}_{12} \\ C_{21} & C_{22} + C_{21}\tilde{S}_n \end{pmatrix} \tilde{v}(t_n). \end{aligned}$$

We choose $\tilde{S}_n$ such that

$$\begin{aligned} \tilde{C}_{12} &=: C_{11}\tilde{S}_n - \tilde{S}_{n+1}C_{21}\tilde{S}_n + C_{12} - \tilde{S}_{n+1}C_{22} \\ &= (C_{11} - I)\tilde{S}_n + C_{12} - \tilde{S}_{n+1}(C_{22} - I) + (\tilde{S}_n - \tilde{S}_{n+1}) + \tilde{S}_{n+1}C_{21}\tilde{S}_n \\ &= \mathcal{O}(k^{p+2}). \end{aligned}$$

We multiply the last equation by $\frac{\varepsilon}{k}(I - \frac{k}{\varepsilon}A_{11})$. Since

$$
\begin{aligned}
\tilde{B}_{ij} &= B_{ij} + kB_{ij}^{(1)},\\
C_{11} - I &= \frac{k}{\varepsilon}(I - \frac{k}{\varepsilon}A_{11})^{-1}(A_{11} + \varepsilon\tilde{B}_{11}),\\
C_{12} &= k(I - \frac{k}{\varepsilon}A_{11})^{-1}\tilde{B}_{12},\\
C_{22} - I &= k(A_{22} + B_{22}) + k^2\tilde{\tilde{B}}_{22},\\
C_{21} &= kB_{21} + k^2\tilde{\tilde{B}}_{21},
\end{aligned}
\tag{4.9}
$$

we obtain

$$
\begin{aligned}
&(A_{11} + \varepsilon B_{11})\tilde{S}_n + \varepsilon\tilde{B}_{12} - (\varepsilon I - kA_{11})\tilde{S}_{n+1}(A_{22} + B_{22} + k\tilde{\tilde{B}}_{22})\\
&\quad + (\varepsilon I - kA_{11})\frac{\tilde{S}_{n+1} - \tilde{S}_n}{k} + (\varepsilon I - kA_{11})\tilde{S}_{n+1}(B_{21} + k\tilde{B}_{21})\tilde{S}_n\\
&= \mathcal{O}(\varepsilon k^{p+1} + k^{p+2}).
\end{aligned}
\tag{4.10}
$$

(4.10) is a $\mathcal{O}(k)$ approximation of (3.8). By iteration, we obtain

Lemma 4.2. *There is a transformation*

$$\tilde{S} = S + \varepsilon kE + \mathcal{O}(k^{p+1})$$

such that, neglecting terms of order $\mathcal{O}(k^{p+1})$.

$$\tilde{v}(t_{n+1}) = \begin{pmatrix} C_{11} - \tilde{S}_{n+1}C_{21} & 0 \\ C_{21} & C_{22} + C_{21}\tilde{S}_n \end{pmatrix} \tilde{v}(t_n). \tag{4.11}$$

As in the previous section, we transform (4.11) to blockdiagonal form by a substitution

$$\tilde{v}(t_n) = \begin{pmatrix} I & 0 \\ \tilde{T}_n & I \end{pmatrix} \tilde{\tilde{v}}(t_n).$$

We obtain

$$\tilde{\tilde{v}}(t_{n+1}) = \begin{pmatrix} C_{11} - \tilde{S}_{n+1}C_{21} & 0 \\ \tilde{C}_{21} & C_{22} + C_{21}\tilde{S}_n \end{pmatrix} \tilde{\tilde{v}}(t_n),$$

where, by (4.9),

$$
\begin{aligned}
\tilde{C}_{21} &= -\tilde{T}_{n+1}(C_{11} - \tilde{S}_{n+1}C_{21}) + C_{21} + C_{22}\tilde{T}_n + C_{21}\tilde{S}_n\tilde{T}_n \\
&= \tilde{T}_{n+1}(I - C_{11}) + C_{21} + (C_{22} - I)\tilde{T}_n - (\tilde{T}_{n+1} - \tilde{T}_n) \\
&\quad + \tilde{T}_{n+1}\tilde{S}_{n+1}C_{21} + C_{21}\tilde{S}_n\tilde{T}_n \\
&= -\frac{k}{\varepsilon}\tilde{T}_{n+1}(I - \frac{k}{\varepsilon}A_{11})^{-1}(A_{11} + \varepsilon\tilde{B}_{11}) + k\tilde{B}_{21} + k^2\tilde{\tilde{B}}_{21} \\
&\quad + (k(A_{22} + \tilde{B}_{22}) + k^2\tilde{\tilde{B}}_{22})\tilde{T}_n - \frac{\tilde{T}_{n+1} - \tilde{T}_n}{k} \\
&\quad + k\tilde{T}_{n+1}\tilde{S}_{n+1}(\tilde{B}_{21} + k\tilde{\tilde{B}}_{21}) + k(\tilde{B}_{21} + k\tilde{\tilde{B}}_{21})\tilde{S}_n\tilde{T}_n.
\end{aligned}
$$

We choose $\tilde{T}$ such that $\tilde{C}_{21} = \mathcal{O}(k^{p+2})$. We divide the above relation by k. Observing that $\frac{1}{\varepsilon}(I - \frac{k}{\varepsilon}A_{11})^{-1} = (\varepsilon I - kA_{11})^{-1}$ and, neglecting all terms of order $\mathcal{O}(k)$, we obtain

$$
\begin{aligned}
&- \tilde{T}_{n+1}(\varepsilon I - kA_{11})^{-1}(A_{11} + \varepsilon B_{11}) + B_{21} + (A_{22} + B_{22})\tilde{T}_{n+1} \\
&+ \frac{\tilde{T}_{n+1} - \tilde{T}_n}{k} + \tilde{T}_{n+1}\tilde{S}_{n+1}B_{21} + B_{21}S_n\tilde{T}_n = \mathcal{O}(k^{p+1}).
\end{aligned} \tag{4.13}
$$

The last relation is a first order accurate approximation of (3.10). We solve it by iteration, starting with

$$
\tilde{T}^{(1)} = B_{21}(A_{11} + \varepsilon B_{11})^{-1}(\varepsilon I - kA_{11}) = T + \mathcal{O}(k).
$$

We have proved

Lemma 4.3. *There is a transformation*

$$
\tilde{T} = T + kE + \mathcal{O}(k^{p+1})
$$

such that , neglecting terms of order $\mathcal{O}(k^{p+1})$,

$$
\tilde{\tilde{v}}(t_{n+1}) = \begin{pmatrix} C_{11} - \tilde{S}_{n+1}C_{21} & 0 \\ 0 & C_{22} + C_{21}\tilde{S}_n \end{pmatrix} \tilde{\tilde{v}}(t_n).
$$

Since

$$|C_{11} - \tilde{S}_{n+1} C_{21}| = \mathcal{O}(\varepsilon/k + k\varepsilon),$$

the first components $\tilde{\tilde{v}}^I(t_n)$ decay rapidly. Also,

$$C_{22} + C_{21}\tilde{S}_n = I + k(A_{22} + B_{22} + B_{21}S) + \mathcal{O}(k^2)$$

and, therefore, by (3.9), there is an expansion

$$\tilde{\tilde{v}}^{II} = \tilde{\tilde{w}} + kE + \mathcal{O}(k^{p+1}).$$

Thus, we have proved Theorem 4.1.

The results in this section show that we can always initialize the data effectively and, therefore, we shall in the following sections only consider slowly varying solutions.

Remark: One could use Richardson extrapolation not only to initialize the data, i.e., use it for a limited number of timesteps, but as a numerical method for all times. However, there are stability problems, especially when A, B do not commute.

5. Backward differentiation methods

Let $k > 0$ denote the time step which defines gridpoints $t_\nu = \nu k$ and gridfunctions $y_\nu = y(\nu k)$. We are interested in the backward differentiation formulas

$$(I - k\alpha_{-1}(\frac{1}{\varepsilon}A + B))y_{n+1} = \sum_{j=0}^{p-1} \alpha_j y_{n-j} \; + k\alpha_{-1}F_{n+1}. \tag{5.1}$$

Here A, B are evaluated at $t_{n+1} = (n+1)k$. Let u be a slowly varying solution of (1.1). Taylor expansion gives us

$$u_{n-j} = u_{n+1} - (j+1)ku'_{n+1} + \frac{(j+1)^2}{2}u''_{n+1} + \dots . \tag{5.2}$$

Therefore,

$$\begin{aligned} L[u] &= (I - k\alpha_{-1}(\frac{1}{\varepsilon}A_{11} + B_{11}))u_{n+1} - \sum_{j=0}^{p-1}\alpha_j u_{n-j} \; - k\alpha_{-1}F_{n+1} \\ &= (1 - \sum_{j=0}^{p-1}\alpha_j)u_{n+1} + k\Big\{(\sum_{j=0}^{p-1}(j+1)\alpha_j)u'_{n+1} - \alpha_{-1}(\frac{1}{\varepsilon}A + B)u_{n+1} - \alpha_{-1}F_{n+1}\Big\} \\ &\quad + \sum_{l=2}^{p} k^l\Big(\sum_{j=0}^{p-1}\frac{(j+1)^l}{l!}\alpha_j\Big)(-1)^l\frac{d^l u_{n+1}}{dt^l} + k^{p+1}R^{(p+1)}_{n+1}, \end{aligned} \tag{5.3}$$

where

$$R^{(p+1)} = c_1\frac{d^{p+1}u}{dt^{p+1}} + kc_2\frac{d^{p+2}u}{dt^{p+2}} + \mathcal{O}(k^2\frac{d^{p+3}u}{dt^{p+3}})$$

with

$$c_\nu = \frac{(-1)^{p+\nu}}{(p+\nu)!}\sum_{j=0}^{p-1}(j+1)^{p+\nu}\alpha_j, \quad \nu = 1, 2.$$

We choose

$$\sum_{j=0}^{p-1}\alpha_j = 1, \quad \sum_{j=0}^{p-1}(j+1)\alpha_j = \alpha_{-1}, \quad \sum_{j=0}^{p-1}(j+1)^l\alpha_j = 0, \quad l = 2, 3, \dots, p. \tag{5.4}$$

Then

$$L[u] = k^{p+1} R^{(p+1)} \tag{5.5}$$

and (5.1) is accurate of order p for slowly varying solutions. The methods are the usual BDF methods and the coefficients can be found in [3].

To obtain useful error estimates the method needs to be stable. For the stability investigation we assume that A, B are constant matrices. In the same way as for the continuous problem, we can separate the slow and the fast scales by transformation. Therefore, we need to investigate stability only for the decoupled systems (2.2a) and (2.2b). The stability question is governed by the stability region of the method when applied to the scalar differential equation

$$y' = \lambda y \tag{5.6}$$

with constant coefficients. We want to use our results also for the solution of hyperbolic partial differential equations, i.e., we want to apply it to systems of ordinary differential equations obtained by discretizing the space operators. As we have shown in [7], a desirable property of the multistep method is that it is locally stable, defined by

Definition 5.1. *A multistep method is called locally stable if the stability region Ω in the complex $\mu = k\lambda$ plane contains a halfdisc*

$$|\mu| = |k\lambda| \le R, \quad \operatorname{Re} k\lambda \le 0.$$

See Figure 1.

For $p = 1$ the scheme (5.1) represents the implicit Euler method which is A-stable and therefore also locally stable. The same is true for $p = 2$.

We shall discuss the stability region for the other values of p in a neighborhood of $\mu = 0$. The solutions of the difference approximation (5.1) are of the form

$$y_n = \kappa^n y_0, \tag{5.7}$$

where κ satisfies

$$(1 - \alpha_{-1}\lambda k)\kappa^p - \sum_{j=0}^{p-1} \alpha_j \kappa^{p-j-1} = 0. \tag{5.8}$$

The basic characteristic equation

$$\kappa^p - \sum_{j=0}^{p-1} \alpha_j \kappa^{p-j-1} = 0 \tag{5.9}$$

has exactly one root $\kappa = 1$ while all other roots satisfy

$$|\kappa_j| < 1, \quad j = 2, 3, \ldots, p.$$

Therefore, in a neighborhood of $\mu = 0$ we need only consider solutions of (5.8) which are of the form

$$\kappa = e^{\lambda k} + k^{p+1}\tau. \tag{5.10}$$

Introducing (5.10) into (5.8) gives us

$$(1 - \alpha_{-1}\lambda k)(e^{\lambda k} + k^{p+1}\tau)^p - \sum_{j=0}^{p-1} \alpha_j (e^{\lambda k} + k^{p+1}\tau)^{p-j-1} = 0.$$

Thus, we obtain the linearized equation

$$\Big(p(1 - \alpha_{-1}\lambda k)(e^{(p-1)\lambda k} - \sum_{j=0}^{p-1} \alpha_j e^{(p-j-2)\lambda k}(p - j - 1)\Big) k^{p+1}\tau = -\tilde{R} \tag{5.11}$$

where

$$\tilde{R} = (1 - \alpha_{-1}\lambda k)e^{\lambda pk} - \sum_{j=0}^{p-1} \alpha_j e^{(p-j-1)\lambda k}.$$

By (5.6) and (5.3), truncation error analysis shows that

$$\begin{aligned} R_{n+1}^{(p+1)} = L[e^{\lambda t}] &= \Big((1 - \alpha_{-1}\lambda k)e^{\lambda pk} - \sum_{j=0}^{p-1} \alpha_j e^{(p-j-1)\lambda k}\Big) e^{(n+1-p)\lambda k} \\ &= \tilde{R} \cdot e^{(n+1-p)\lambda k}. \end{aligned}$$

Therefore, from (5.11) we obtain, using Taylor expansion and (5.4),

$$(5.12) \qquad (\alpha_{-1} + d_0 \lambda k + \mathcal{O}((\lambda k)^2))\tau = c_1 \lambda^{p+1} + d_1 \lambda^{p+2} k + \mathcal{O}(\lambda^{p+3} k^2).$$

Here d_0, d_1 are real constants which can easily be calculated. The last relation finally gives us

$$\begin{aligned} \kappa &= e^{\lambda k} + k^{p+1}\tau \\ &= e^{\lambda k} + (\lambda k)^{p+1} \frac{c_1 + d_1 \lambda k}{\alpha_{-1} + d_0 \lambda k} + \mathcal{O}((\lambda k)^{p+3}) \\ &=: e^{\lambda k} + a_1(\lambda k)^{p+1} + a_2(\lambda k)^{p+2} + \mathcal{O}((\lambda k)^{p+3}). \end{aligned}$$

We now assume that $\lambda k = i\mu$ is purely imaginary. Then

$$(5.13) \qquad |\kappa|^2 = \begin{cases} 1 + 2a_1(-1)^{\frac{p+1}{2}} |\mu|^{p+1} + \mathcal{O}(|\mu|^{p+2}) & \text{if } p \equiv 1(2), \\ 1 - 2(a_1 - a_2)(-1)^{\frac{p+2}{2}} |\mu|^{p+2} + \mathcal{O}(|\mu|^{p+3}) & \text{if } p \equiv 0(2). \end{cases}$$

Thus, an interval on the imaginary axis belongs to Ω if

$$(5.13a) \qquad \begin{aligned} (-1)^{\frac{p+1}{2}} a_1 < 0 \quad &\text{for} \quad p \equiv 1(2), \\ (-1)^{\frac{p+2}{2}} (a_1 - a_2) > 0 \quad &\text{for} \quad p \equiv 0(2). \end{aligned}$$

In this case the method is locally stable. Simple calculations show that the method is not locally stable for $p = 3, 4$ but it is locally stable for $p = 5, 6$. See, for example, [3].

For $p = 3, 4, 5, 6$ we shall discuss another property of the stability region. Assume that

$$(5.14) \qquad \lambda k = i\mu - \tau\mu^l, \quad l = 2[\frac{p+1}{2}], \quad ([x] \text{ largest integer } \leq x),$$

where $\tau > 0$ is a constant which we choose later. In this case (5.13) becomes

$$|\kappa|^2 = \begin{cases} 1 - 2(\tau - a_1(-1)^{\frac{p+1}{2}})\mu^{p+1} + \mathcal{O}(\mu^{p+2}) & \text{if } p \equiv 1(2), \\ 1 - 2(\tau + (a_1 - a_2)(-1)^{\frac{p+2}{2}})\mu^{p+2} + \mathcal{O}(\mu^{p+3}) & \text{if } p \equiv 0(2). \end{cases} \tag{5.15}$$

Thus, the stability region Ω always contains a region $\Omega_0 \cap \Omega_1$, $\Omega_1 \cap \Omega_0$, where

$$\begin{aligned} &\Omega_0 : \{\text{Re } \lambda k \leq 0, \quad |\text{Re } \lambda k| \geq \tau |\text{Im } \lambda k|^l, \quad l = 2[\frac{p+1}{2}], \\ &\Omega_1 : \{|\lambda k| \leq R, \end{aligned} \tag{5.16}$$

provided $R > 0$ is sufficiently small and τ sufficiently large. ($\tau = 0$ for $p = 5, 6$.)

The usual calculations show that, for sufficiently large τ, the whole region $\Omega_0 \in \Omega$. In Table 5.1 the minimum values of τ are calculated.

p	3	4	5	6
τ	0.22	0.3	0.3	190

Table 5.1

It shows that, for $p = 6$, τ has to be very large and the method is not useful.

We want to express the relation (5.16) also in another way. We assume that $A + B$ is blockdiagonal and consider the slow part

$$(I - k(A_{22} + B_{22}))y_{n+1}^{II} = \sum_{j=0}^{p-1} \alpha_j y_{n-j}^{II}. \tag{5.17}$$

We shall prove

Theorem 5.1. *Consider (5.17) for sufficiently small $k(|A_{22}| + |B_{22}|)$. We can stabilize (5.17) by changing $A_{22} + B_{22}$ to $A_{22} + B_{22} - \sigma(k(|A_{22}| + |B_{22}|))^l I$, $l = 2[(p+1)/2]$, provided σ is sufficiently large.*

Proof. The eigenvalues λk are changed to $\lambda k - \sigma\big(k(|A_{22}| + |B_{22}|)\big)^{l}$. Observing that $|\mathrm{Im}\ \lambda k| \leq k(|A_{22}| + |B_{22}|)$ the theorem follows from (5.16).

6. Applications

Consider the scalar wave equation

$$(6.1)\qquad \begin{aligned} w_t &= aw_x, \quad 0 \le x \le 2\pi, \\ w(x,0) &= f, \end{aligned}$$

with periodic initial and boundary conditions. We discretize space by introducing gridpoints $x_\nu = \nu h$, $\nu = 0, \pm 1, \pm 2, \ldots$; $h = 2\pi/N$, and gridfunctions $u_\nu(t) = u(x_\nu, t)$ and approximate (6.1) by the dissipative approximation

$$(6.2)\qquad \begin{aligned} u_t &= aD_0 u - \sigma h^3 D_+^2 D_-^2 u, \quad u = u_\nu, \\ u(0) &= f. \end{aligned}$$

Here D_0, D_+, D_- are the usual centered, forward and backward difference operators, respectively. After Fourier transform (6.2) becomes

$$(6.3)\qquad \hat{u}_t = (a\frac{i \sin \omega h}{h} - \frac{16\sigma}{h}\sin^4\frac{\omega h}{2})\hat{u} =: \hat{Q}\hat{u}.$$

For time integration we choose the BDF methods. For $p = 1, 2$, the approximations are unconditionally stable. For $p = 3, 4$, we have to choose the timestep k and the dissipative coefficient σ such that $\hat{Q}$ belongs to the stability region. By (4.16) and Table 5.1, we need that

$$k\frac{16\sigma}{h}\sin^4\frac{\omega h}{2} \ge \frac{\tau k^4 \sin^4 \omega h}{h^4} a^4,$$

for all ωh, i.e.,

$$(6.4)\qquad \sigma\frac{k}{h} \ge \tau(\frac{ka}{h})^4 \quad \text{or} \quad \sigma \ge \tau|a|(\frac{k|a|}{h})^3.$$

Thus, the amount of dissipation becomes rather large if we want to beat the speed of propagation, i.e., choose the timestep k such that $k|a|/h >> 1$.

For $p = 5, 6$, we can also choose τ such that (6.4) holds. For $p = 6$, we need to choose τ very large and, therefore, the method is not useful. Since for $p = 5, 6$, the schemes are locally stable, we can use $\tau = 0$, but then

$$|a\frac{k}{h}| \leq R. \tag{6.5}$$

The restriction (6.5) is more severe than the stability restriction for a good explicit method and, therefore, we shall not pursue this possibility further.

Now consider a hyperbolic system

$$\begin{aligned} w_t &= Aw_x, \quad 0 \leq x \leq 2\pi, \\ w(x,0) &= f, \end{aligned} \tag{6.6}$$

and approximate it by

$$u_t = AD_0 u - \sigma h^3 D_+ D_- u. \tag{6.7}$$

Let $a_1, \ldots, a_n$ denote the eigenvalues of A. By (6.4) we need to choose

$$\sigma \geq \tau \max_j |a_j| \Big(\frac{k \max_j |a_j|}{h}\Big)^3. \tag{6.8}$$

If all the a_j are of the same order of magnitude, this is not a severe restriction. However, for problems with different time scales, they are of different orders of magnitude. Then (6.7) is useless because the dissipation becomes too large. We can replace (6.7) by

$$u_t = AD_0 u - \sigma A^4 h^3 D_+ D_- u.$$

We transform the system to diagonal form and obtain

$$u_t = aD_0 u - \sigma a^4 h^3 D_+ D_- u, \quad a = a_j.$$

Our condition (6.8) now becomes

$$\sigma \geq \tau(\frac{k}{h})^3. \tag{6.9}$$

(6.9) is satisfactory. However, it is expensive to calculate the dissipation term. This is particularly true for problems in more than one space dimension. Therefore, we believe that only the first and second order BDF methods are useful to solve hyperbolic partial differential equations with different time scales.

7. Examples of semi-implicit methods

In many applications it is much easier to calculate $(I-\alpha_{-1}\frac{k}{\varepsilon}A)^{-1}$ than $(I-k\alpha_{-1}(\frac{1}{\varepsilon}A+B))^{-1}$. We can still use the backward differentiation formulas but we solve them by iteration.

$$(7.1)\qquad (I-\alpha_{-1}\frac{k}{\varepsilon}A_{n+1})y_{n+1}^{[j+1]}=\sum_{j=0}^{p-1}\alpha_j y_{n-j}+\alpha_{-1}kB_{n+1}y_{n+1}^{[j]},\quad y^{[0]}\equiv 0.$$

(For simplicity only, we assume that $F\equiv 0$.)

Now assume that $A=-A^*$, $B=-B^*$ are skew-Hermitean. Without restriction we can also assume that A is blockdiagonal.

We want to investigate how many iterations we need to make such that the resulting method is stable. As we have seen in the previous section, only the first and second order BDF methods are useful for our purposes. We consider the second order method. Since there are no stability problems with the large part, we need only consider

$$(I-\frac{2k}{3}A)y_{n+1}^{[j+1]}=\frac{4}{3}y_n-\frac{1}{3}y_{n-1}+\frac{2}{3}kBy_{n+1}^{[j]},\quad y^{[0]}\equiv 0.$$

Here we have deleted the index 22. We stop after s iterations and obtain

$$\begin{aligned}
y_{n+1}&=(I-\frac{2k}{3}A)^{-1}\sum_{j=0}^{s-1}(\frac{2k}{3}B(I-\frac{2k}{3}A)^{-1})^j(\frac{4}{3}y_n-\frac{1}{3}y_{n-1})\\
&=(I-\frac{2k}{3}A)^{-1}(I-\frac{2k}{3}B(I-\frac{2k}{3}A)^{-1})^{-1}\Big(I-(\frac{2k}{3}B(I-\frac{2k}{3}A)^{-1})^s\Big)(\frac{4}{3}y_n-\frac{1}{3}y_{n-1})\\
&=(I-\frac{2k}{3}(A+B))^{-1}\Big(I-(\frac{2k}{3}B)^s+\mathcal{O}(k^{s+1}|B|^s(|B|+|A|))\Big)(\frac{4}{3}y_n-\frac{1}{3}y_{n-1}).
\end{aligned}$$

The worst stability situation occurs when $A=-B$. Then, to first approximation,

$$y_{n+1}=(I-(\frac{2k}{3}B)^s)(\frac{4}{3}y_n-\frac{1}{3}y_{n-1}).$$

Clearly, the approximation is not stable for $s = 1, 2$. For $s > 2$, we can write the approximation as

$$(I - kC)y_{n+1} = \frac{4}{3}y_n - \frac{1}{3}y_{n-1},$$

where

$$kC = \frac{2k}{3}(A+B) + (\frac{2k}{3}B)^s + \mathcal{O}(k^{s+1}|B|^s(|B|+|A|)).$$

The approximation is stable if $C + C^* \geq 0$. For $s \geq 3$ and $k(|A| + |B|)$ sufficiently small, we can, corresponding to the previous section, always stabilize the method by changing $A + B$ to $A + B - \sigma k^4 I$. Here $\sigma =$ const. $(|A^4| + |B^4|)$. If $s = 4$, we do not need the extra stabilization if B is nonsingular. The disadvantage of the method is that we need to invert the operator $I - \frac{2k}{3}A$ at least three times.

Another approach is to split the operator and use a mixed implicit-explicit approximation. The classic method is a combination of the Leap-frog and Crank-Nicholson methods.

$$(I - \frac{k}{\varepsilon}A_{n+1})y_{n+1} = (I + \frac{k}{\varepsilon}A_{n-1})y_{n-1} + 2kB_n y_n. \tag{7.2}$$

It has been used for a long time and is second order accurate for slowly varying solutions and stable if $A + A^* \leq 0$, $B = -B^*$, $|Bk| \leq 1$. As in the previous section, one needs a small timestep to calculate the solutions on the fast scale.

Let us discuss the behavior of the method with respect to the fast scale when $k >> \varepsilon$. Without restriction we can assume that A already has the blockdiagonal form (1.3). Then, neglecting terms of order $\mathcal{O}(\varepsilon)$, the equation for the fast scale becomes

$$y^I_{n+1} = (I - \frac{k}{\varepsilon}A_{11})^{-1}(I + \frac{k}{\varepsilon}A_{11})y^I_{n-1},$$

and, therefore, for $k >> \varepsilon$,

$$y^I_{n+1} \simeq -y^I_{n-1}. \tag{7.3}$$

Thus, the fast scale behaves like a ± 1-wave and will not be rapidly damped.

If we have initialized the data properly, then the fast scale solution is not present and waves are not excited. However, if the initialization has not been done carefully or if fast waves are generated by, for example, rough data, then we will have difficulties. The remedy is to use time filters (see, for example, [8]) or use the Richardson extrapolation of Section 4 to re-initialize the solution periodically. To avoid this problem altogether one often uses a combination of the backward Euler and Leap-frog methods. Again, assuming that A is blockdiagonal, the scheme is given by

$$(I - \frac{2k}{\varepsilon}A_{11})y^I_{n+1} = y^I_{n-1} + 2k(B_{11}y^I_n + B_{12}y^{II}_n), \tag{7.4a}$$

$$(I - 2kA_{22})y^{II}_{n+1} = y^{II}_{n-1} + 2k(B_{21}y^I_n + B_{22}y^{II}_n). \tag{7.4b}$$

As in Section 4, we can derive an asymptotic expansion and obtain, for y^I,

$$y^I_n = \tilde{y}^I_n + \varphi^I_0(t_n), \tag{7.5a}$$

where

$$\varphi^I_0(t_{n+1}) = -\varepsilon A_{11}^{-1}B_{12}y^{II}_n + \mathcal{O}(\varepsilon^2),$$

$$(I - \frac{2k}{\varepsilon}A_{11})\tilde{y}^I_{n+1} = \tilde{y}^I_{n-1}.$$

Now the rapidly varying part is decaying quickly. Introducing (7.5a) into (7.4b), neglecting terms of order $\mathcal{O}(k\varepsilon^2)$, gives us

$$(I - 2kA_{22})y^{II}_{n+1} = (I + 2k\varepsilon\tilde{B}_{22})y^{II}_{n-1} + 2k(B_{21}\tilde{y}^I_n + B_{22}y^{II}_n), \tag{7.5b}$$

where

$$\tilde{B}_{22} = -B_{21}A_{11}^{-1}B_{12}.$$

Now we will discuss the accuracy and the stability of (7.5b) for slowly varying solutions ($y_n^I \equiv 0$). If $A_{22} \neq 0$, then (7.5b) is only a first order approximation of (1.1). Also, we cannot expect it to be stable. To show this we apply it to a scalar problem:

$$A_{22} = ia, \quad B_{21} = 0, \quad B_{22} = ib,$$

i.e.,

$$(1 - 2ika)y_{n+1} = y_{n-1} + 2ikby_n.$$

We construct solutions of the form

$$y_n = \kappa^n y_0,$$

where κ are the solutions of the characteristic equation

$$(1 - 2ika)\kappa^2 - 2ikb\kappa - 1 = 0,$$

i.e.,

$$\begin{aligned}\kappa &= \frac{1}{1-2ika}(kbi \pm \sqrt{1 - 2ika - k^2b^2}) \\ &= \frac{1}{1-2ika}\Big(ikb \pm (1 - ika - \frac{k^2b^2}{2} + \frac{k^2a^2}{2} + \mathcal{O}(k^3))\Big).\end{aligned}$$

Therefore,

$$\begin{aligned}|\kappa^2| &= \frac{1}{1+4k^2a^2}\Big(1 + k^2(b \mp a)^2 - k^2b^2 + k^2a^2\Big) + \mathcal{O}(k^4) \\ &= \frac{1 \mp 2k^2ab + 2k^2a^2}{1+4k^2a^2} + \mathcal{O}(k^4).\end{aligned}$$

If $ab \neq 0$ and $|b| > |a|$, then there is a k_0 such that $|\kappa| > 1$ for $0 < k < k_0$. Thus, the method is not locally stable.

Remark: One can prove the following theorem which is important for reaction-diffusion equations.

Theorem 7.1. *Assume that* $A_{11} = A_{11}^* \le 0$, $\tilde{B}_{22} = \tilde{B}_{22}^* \le 0$, $B_{22} = -B_{22}^*$. *Then the method (7.5b) is stable if* $k(|B_{22}| + \varepsilon|\tilde{B}_{22}|) < 1$.

If $A_{22} \equiv 0$, then the method is, formally, still first order, but since $\varepsilon\tilde{B}_{22}$ is of order $\mathcal{O}(\varepsilon)$, the error is of order $\mathcal{O}(\varepsilon k + k^2)$. By assumption, $\varepsilon << k$ and, therefore, it is a second order method.

In order to investigate the stability problems of (7.5b) we consider the scalar equation

$$(7.6) \qquad y_{n+1} = (1 + 2k\varepsilon\tilde{b})y_{n-1} + 2kiby_n, \quad \tilde{b} = \tilde{b}_1 + i\tilde{b}_2, \quad b_1, b_2 \text{ real.}$$

The characteristic equation now becomes

$$\kappa^2 - 2kib\kappa - (1 + 2k\varepsilon\tilde{b}) = 0.$$

Therefore, for $1 - k^2b^2 >> |k\varepsilon\tilde{b}|$,

$$\begin{aligned}\kappa &= kib \pm \sqrt{1 - k^2b^2 + 2k\varepsilon\tilde{b}} \\ &= kib \pm \left(\sqrt{1 - k^2b^2} + \frac{k\varepsilon\tilde{b}}{\sqrt{1 - k^2b^2}} + \mathcal{O}(\varepsilon^2k^2)\right),\end{aligned}$$

i.e.,

$$\begin{aligned}|\kappa|^2 &= (kb \pm k\varepsilon\tilde{b}_2)^2 + 1 - k^2b^2 + 2k\varepsilon\tilde{b}_1 + \mathcal{O}(\varepsilon k^4 + \varepsilon^2k^2) \\ &= 1 \pm 2k^2\varepsilon b\tilde{b}_2 + 2k\varepsilon\tilde{b}_1 + \mathcal{O}(\varepsilon k^4 + \varepsilon^2k^2).\end{aligned}$$

Thus, neglecting terms of order $\mathcal{O}(\varepsilon k^4 + \varepsilon^2 k^2)$,

$$|\kappa| \le 1 \quad \text{if} \quad \tilde{b}_1 \le -k|b\,b_2|.$$

(See Theorem 7.1.) If $\tilde{b}_1 = 0$, and $|b\,b_2| \ne 0$, then there is growth. However, $|\kappa|^2 = 1 + k^2\mathcal{O}(\varepsilon)$ and, therefore, the amplification is weak and does not destroy the calculations in time intervals of length $\mathcal{O}(1/\varepsilon)$. Similar results hold for systems.

8. General multistep methods

In this section we consider general approximations of the form

$$(8.1) \qquad (I - \alpha_{-1}\frac{k}{\varepsilon}A_{n+1})y_{n+1} = \sum_{j=0}^{r-1}\alpha_j y_{n-j} + k\sum_{j=0}^{q-1}\beta_j B_{n-j}y_{n-j}.$$

(For simplicity only, we assume that $F \equiv 0$.) We are interested in methods which are accurate of order p and we first look at the case when $B \equiv 0$. We determine the coefficients α_j such that

$$\begin{aligned} u_{n+1} - k\alpha_{-1}\frac{du_{n+1}}{dt} &- \sum_{j=0}^{r-1}\alpha_j u_{n-j} \\ &= c_{p+1}\frac{k^{p+1}}{(p+1)!}\frac{d^{p+1}u_{n+1}}{dt^{p+1}} + c_{p+2}\frac{k^{p+2}}{(p+2)!}\frac{d^{p+2}u_{n+1}}{dt^{p+2}} + \mathcal{O}(k^{p+3}), \end{aligned}$$

for all smooth functions. Using Taylor expansion, this leads to the linear equations

$$(8.2) \qquad \sum_{j=0}^{r-1}\alpha_j = 1, \quad \sum_{j=0}^{r-1}(j+1)\alpha_j = \alpha_{-1}, \quad \sum_{j=0}^{r-1}(j+1)^l\alpha_j = 0, \quad l = 2,3,\ldots,p\,.$$

Also,

$$(8.3) \qquad c_\nu = -(-1)^\nu\sum_{j=0}^{r-1}(j+1)^\nu\alpha_j.$$

If we set $r = p$ we recover the BDF formulas (4.2).

Now assume that we have chosen α_j such that (8.2) holds. We will apply the method for the case that $A \equiv 0$. We want to choose the β_j such that the resulting method

$$(8.4) \qquad y_{n+1} = \sum_{j=0}^{r-1}\alpha_j y_{n-j} + k\sum_{j=0}^{q-1}\beta_j B_{n-j}y_{n-j}$$

is accurate of order s with $p \le s \le q$. In this case $By = dy/dt$ and, therefore, we have to choose β_j such that, for all functions u,

$$(8.5) \qquad u_{n+1} - \sum_{j=0}^{r-1}\alpha_j u_{n-j} - k\sum_{j=0}^{q-1}\beta_j\frac{du_{n-j}}{dt} = \mathcal{O}(k^{s+1}).$$

Using (8.2), we have,

$$u_{n+1} - \sum_{j=0}^{r-1} \alpha_j u_{n-j} - k \sum_{j=0}^{q-1} \beta_j \frac{du_{n-j}}{dt}$$

$$= k\alpha_{-1} \frac{d_{n+1}}{dt} - \sum_{l=p+1}^{s} \frac{(-1)^l k^l}{l!} \Big(\sum_{j=0}^{r-1} (j+1)^l \alpha_j\Big) \frac{d^l u_{n+1}}{dt^{l+1}} - k \Big(\sum_{j=0}^{q-1} \beta_j\Big) \frac{du_{n+1}}{dt}$$

$$- \sum_{l=2}^{s} \frac{(-1)^l k^l}{(l-1)!} \sum_{j=0}^{q-1} (j+1)^{l-1} \beta_j \frac{d^l u_{n+1}}{dt^l} + \mathcal{O}(k^{s+1}).$$

Therefore, (8.5) is satisfied if

$$\text{(8.6)} \qquad \begin{aligned} &\sum_{j=0}^{q-1} \beta_j = \alpha_{-1}, \quad \sum_{j=0}^{q-1} (j+1)^{l-1} \beta_j = 0, \quad l = 2, \ldots, p\,, \\ &\frac{1}{l} \sum_{j=0}^{r-1} (j+1)^l \alpha_j = \sum_{j=0}^{q-1} (j+1)^{l-1} \beta_j, \quad l = p+1, \ldots, s\,. \end{aligned}$$

If the conditions (8.3),(8.6) are satisfied, then we obtain, for smooth functions,

$$\text{(8.7)} \qquad \begin{aligned} L[u] =: (I - \alpha_{-1} \frac{k}{\varepsilon} A_{n+1}) u_{n+1} - \sum_{j=0}^{r-1} \alpha_j u_{n-j} - k \sum_{j=0}^{q-1} \beta_j B_{n-j} u_{n-j} \\ = k\alpha_{-1} \Big(\frac{du_{n+1}}{dt} - (\frac{1}{\varepsilon} A_{n+1} + B_{n+1}) u_{n+1} \Big) \\ + k^{p+1} R_{n+1}^{(p+1)} + k^{p+2} R_{n+1}^{(p+2)} + \mathcal{O}(k^{p+3}), \end{aligned}$$

where

$$\text{(8.8)} \qquad R^{(\nu)} = \frac{c_\nu}{\nu!} \frac{d^\nu u}{dt^\nu} + \frac{d_{\nu-1}}{(\nu-1)!} \frac{d^{\nu-1}(Bu)}{dt^{\nu-1}}.$$

Corresponding to (8.3),

$$\text{(8.9)} \qquad d_\nu = -(-1)^\nu \sum_{j=0}^{q-1} (j+1)^\nu \beta_j$$

Thus, for slowly varying solutions, the error is at most $\mathcal{O}(k^{p+1})$.

3

As we have seen in the previous sections, the most interesting methods are obtained when $r = p = 2$ and $s = q = 2, 3, 4, 5, 6$. In this case

$$\alpha_{-1} = \frac{2}{3}, \quad \alpha_0 = \frac{4}{3}, \quad \alpha_1 = -\frac{1}{3}.$$

We have calculated the β_j in Table 8.1.

q	β_0	β_1	β_2	β_3	β_4	β_5
2	1.3333	−0.6667				
3	1.7778	−1.5556	0.4444			
4	2.1667	−2.7222	1.6111	−0.3889		
5	2.5241	−4.1519	3.7556	−1.8185	0.3574	
6	2.8602	−5.8324	7.1167	−5.1796	2.0380	−0.3361

Table 8.1

We will now discuss the stability of

$$(I - \frac{2}{3}\frac{k}{\varepsilon}A)u_{n+1} = \frac{4}{3}u_n - \frac{1}{3}u_{n-1} + kB\sum_{j=0}^{q-1}\beta_j u_{n-j}. \tag{8.10}$$

We assume that A, B are constant matrices.

If $A \equiv 0$, then the stability question is reduced to the usual scalar problem, i.e., the behavior of the roots of the characteristic equation

$$\kappa^q - \sum_{j=0}^{q-1}(\alpha_j + \lambda k\beta_j)\kappa^{q-j-1} = 0, \quad \alpha_j = 0 \text{ for } j > 1.$$

Since the basic characteristic equation ($\lambda k = 0$) is the same as for the BDF methods, we need only consider the perturbation

$$\kappa = e^{\lambda k} + k^{p+1}\tau.$$

For τ, $|\kappa|^2$ we obtain relations of type (5.12) and (5.13), respectively. For $q = 3,4$, the methods are locally stable and (5.13a) holds. For $q = 2,5,6$, they are not locally stable. In Figure 2 we have calculated part of the stability region.

Now let $A \neq 0$. Without restriction we can assume that A has the blockdiagonal form (1.3). Then the first components ($\alpha_0 = 4/3$, $\alpha_1 = -1/3$, $\alpha_j = 0$ for $j > 1$) of the system satisfy

$$(8.11) \qquad (I - \frac{2k}{3\varepsilon}A_{11})y^I_{n+1} = \sum_{j=0}^{q-1}\Big((\alpha_j I + k\beta_j B)y_{n-j}\Big)^I.$$

As in the previous section, neglecting terms of order $\mathcal{O}(\varepsilon^2)$, the solution of (8.11) is

$$(8.12a) \qquad y^I_{n+1} = -\varepsilon\sum_{j=0}^{q-1}\beta_j A_{11}^{-1}B_{12}y^{II}_{n-j} + \tilde{y}^I,$$

$$(8.12b) \qquad (I - \frac{2k}{3\varepsilon}A_{11})\tilde{y}^I_{n+1} = \sum_{j=0}^{q-1}\alpha_j\tilde{y}^I_{n-j}.$$

(8.12b) is the second order BDF method which is unconditionally stable. Neglecting terms of order $\mathcal{O}(\varepsilon)$, we obtain, for the other components,

$$(8.13) \qquad (I - \frac{2k}{3}A_{22})\tilde{y}^{II}_{n+1} = \sum_{j=0}^{q-1}(\alpha_j I + k\beta_j B_{22})y^{II}_{n-j}.$$

If $A_{22} = 0$, the above stability results apply. If $A_{22} \neq 0$ and A_{22} does not commute with B_{22}, we cannot reduce the stability question to a scalar equation. Instead, we must consider solutions of (8.13) which are of the form

$$(8.14) \qquad y_n = \kappa^n y_0, \quad y_o \text{ vector},$$

where

$$\kappa = e^{\lambda k} + (\lambda k)^3 \tau, \quad y_0 = u_0 + (\lambda k)^2 y_1,$$

and $u = e^{\lambda t} u_0$ is the corresponding solution of

$$u_t = (A_{22} + B_{22})u,$$

i.e.,

$$(A_{22} + B_{22})u_o = \lambda u_0. \tag{8.15}$$

Let

$$(\lambda k)^3 \tilde{R} = \Big((I - \frac{2k}{3} A_{22}) e^{\lambda q k} - \sum_{j=0}^{q-1} (\alpha_j I + k\beta_j B_{22}) e^{\lambda(q-j-1)k}\Big) u_0.$$

Introducing (8.14) into (8.13) and deleting the index 22, we obtain, using (8.2) and (8.6),

$$\begin{aligned}
0 &= \Big((I - \frac{2k}{3} A)(e^{\lambda k} + (\lambda k)^3 \tau)^q\Big) \\
&\quad - \Big(\sum_{j=0}^{q-1} (\alpha_j I + k\beta_j B)(e^{\lambda k} + (\lambda k)^3 \tau)^{q-1-j}\Big)(u_0 + (\lambda k)^2 y_1) \\
&= (\lambda k)^{p+1} \tilde{R} + (\lambda k)^3 \tau \Big(q - \sum_{j=0}^{q-1} \alpha_j (q-j-1) \\
&\quad + \lambda k \big(q(q-1) - \sum_{j=0}^{q-1} \alpha_j (q-(j+1))(q-1-(j+1))\big)\Big) u_0 \\
&\quad - (\lambda k)^3 \tau \Big(\frac{2}{3} qkA + kB \sum_{j=0}^{q-1} \beta_j (q-j-1)\Big) u_0 \\
&\quad + (\lambda k)^2 \Big(\lambda k q + \frac{1}{2}\lambda^2 q^2 k^2 - \lambda k \sum_{j=0}^{q-1} \alpha_j (q-j-1) - \frac{1}{2}\lambda^2 k^2 \sum_{j=0}^{q-1} \alpha_j (q-j-1)^2\Big) y_1 \\
&\quad - (\lambda k)^2 k \Big(\frac{2}{3} A(1 + \lambda q k) + B \sum_{j=0}^{q-1} \beta_j \big(1 + \lambda k (q-j-1)\big)\Big) y_1 + \mathcal{O}(\lambda k)^5 \\
&= (\lambda k)^3 \tilde{R} + \frac{2}{3}(\lambda k)^3 \tau (1 + (2q-1)\lambda k) u_0 - \frac{2}{3}(\lambda k)^3 \tau q k (A+B) u_0 \\
&\quad + \frac{2}{3}(\lambda k)^3 (1 + \lambda q k) y_1 - \frac{2}{3}(\lambda k)^2 k (1 + \lambda q k)(A+B) y_1 + \mathcal{O}\big((\lambda k)^5\big).
\end{aligned}$$

Therefore, (8.15) gives us,

$$
\begin{aligned}
&\tau(1-(q-1)\lambda k)u_0 + \frac{1}{\lambda}(1+\lambda qk)(\lambda I-(A+B))y_1 \\
&= -\frac{3}{2}\tilde{R} + \mathcal{O}(k^2).
\end{aligned}
\tag{8.16}
$$

We can express $\tilde{R}$ in terms of the truncation error. By (8.7), replacing $\frac{1}{\varepsilon}A_{n+1}$ by A and B_{n-j} by B we obtain, using (8.15),

$$
\begin{aligned}
(\lambda k)^3\tilde{R} &= L[e^{\lambda t}u_0]e^{-\lambda(n+1-q)k} \\
&= \left(k^3R^{(3)}_{n+1} + k^4R^{(4)}_{n+1}\right)e^{-(n+1-q)k} + \mathcal{O}(k^5) \\
&= \frac{(\lambda k)^3}{3!}e^{\lambda qk}\left(c_3I + 3d_2\frac{1}{\lambda}B\right)u_0 + \frac{(\lambda k)^4}{4!}\left(c_4I + 4d_3\frac{1}{\lambda}B\right)u_0 \\
&= \frac{(\lambda k)^3}{3!}e^{\lambda qk}\left((c_3+3d_2)I - 3d_2\frac{1}{\lambda}A\right)u_0 \\
&\quad + \frac{(\lambda k)^4}{4!}\left((c_4+4d_3)I - 4d_3\frac{1}{\lambda}A\right)u_0 + \mathcal{O}(k^5).
\end{aligned}
\tag{8.17}
$$

We now make

Assumption 8.1. $A = -A^*$, $B = -B^*$ *are skew-Hermitean.*

Normalizing the eigenvector u_0 such that $|u_0| = 1$, we obtain, from (8.16) and (8.17),

$$
\tau = (\lambda k)^3(a_1 + a_2\langle u_0, \frac{1}{\lambda}Au_0\rangle) + (\lambda k)^4(b_1 + b_2\langle u_0, \frac{1}{\lambda}Au_0\rangle) + \mathcal{O}(k^5).
$$

Here the real constants a_j, b_j, $j = 1, 2$, can easily be calculated from the previous expressions. Therefore, $|\kappa|^2$ is of the form (5.13) and, corresponding to Theorem 5.1, we have

Theorem 8.1. *For sufficiently small* $k(|A_{22}| + |B_{22}|)$, *the methods*

$$
(I - \frac{2}{3}\frac{k}{\varepsilon}A)y_{n+1} = \sum_{j=0}^{q-1}(\alpha_jI + k\beta_jB)y_{n-j}
$$

are stabilized if we replace B *by* $B - \sigma\big(k(|A_{22}| + |B_{22}|)\big)^4 I$, *provided* σ *is sufficiently large.*

Since, for $q = 3$ and $A = 0$, the method is locally stable, we can use $\sigma = 0$ if $\langle \hat{u}_0, \frac{1}{\lambda} A_{22} \hat{u}_0 \rangle$ is sufficiently small. As an example, we have considered the differential equation

$$y' = (ig|c| + c)y, \quad c = c_1 + ic_2, \quad c_1, c_2 \text{ real}, \ c_1 \leq 0,$$

and approximated it by the combination of second order BDF with the third order explicit method

$$(8.18) \qquad (1 - \frac{2k}{3} g|c|) y_{n+1} = \sum_{j=0}^{3} (\alpha_j I + k\beta_j c) y_{n-j}.$$

In Figure 3 we have calculated part of the stability region for different values of g.

REFERENCES

1. G. Browning and H.O. Kreiss," Splitting methods for problems with different time scales," *Monthly Weather Rev.*, (1994) Vol.**122** pp.2614-2622.
2. B. Engquist, E. Haas and H.O. Kreiss, "Numerical solution of problems with different time scales I", To appear.
3. C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations.* Prentice-Hall, New Jersey, 1971.
4. H.O. Kreiss, "Problems with different time scales for ordinary differential equations", *Siam J. Numer. Anal.*, (1979) Vol.**6** No.6 (1979).
5. H.O. Kreiss, "Problems with different time scales", *Acta Numerica*, (1991), pp.1-39.
6. H.O. Kreiss and J. Lorenz, "On the existence of slow manifolds for problems with different time scales", *Phil. Trans. R. Soc. Lond.*, (1994) Vol.**346**, pp.159-171.
7. H.O. Kreiss and L. Wu, "On the stability definition of difference approximations for the initial boundary value problem", *Appl. Num. Math.* (1993) Vol.**12** pp.213-227.
8. B. Lindberg, "On the smoothing and extrapolation for the trapezoidal rule",*Bit* (1971) Vol.**11** pp.29-52.

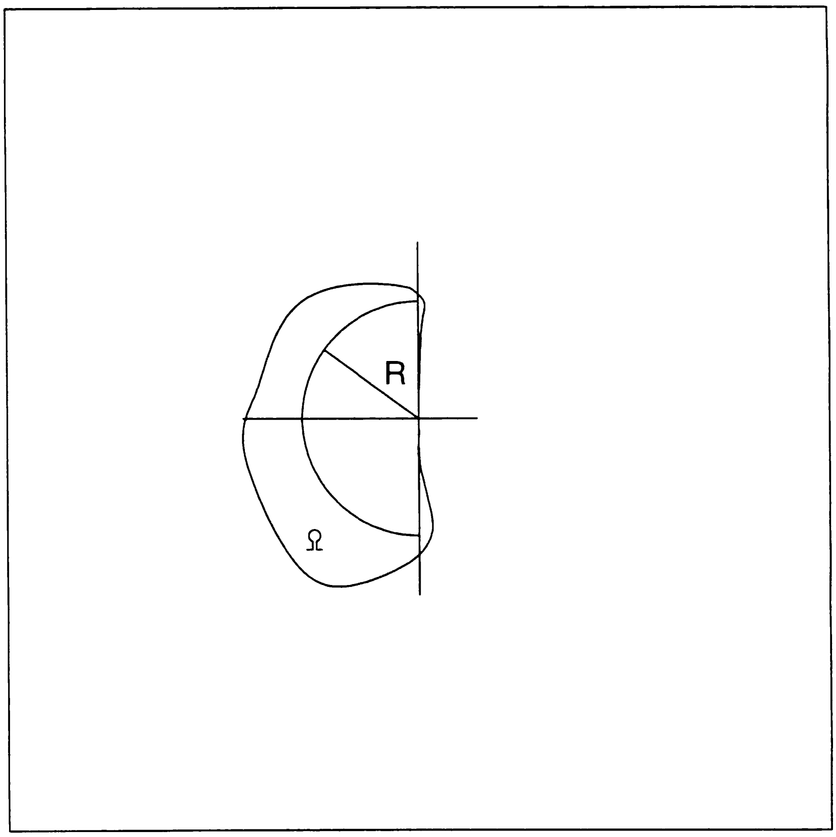

Figure 1

Stability region for (8.10) where A=0, q=2,3,4,5,6

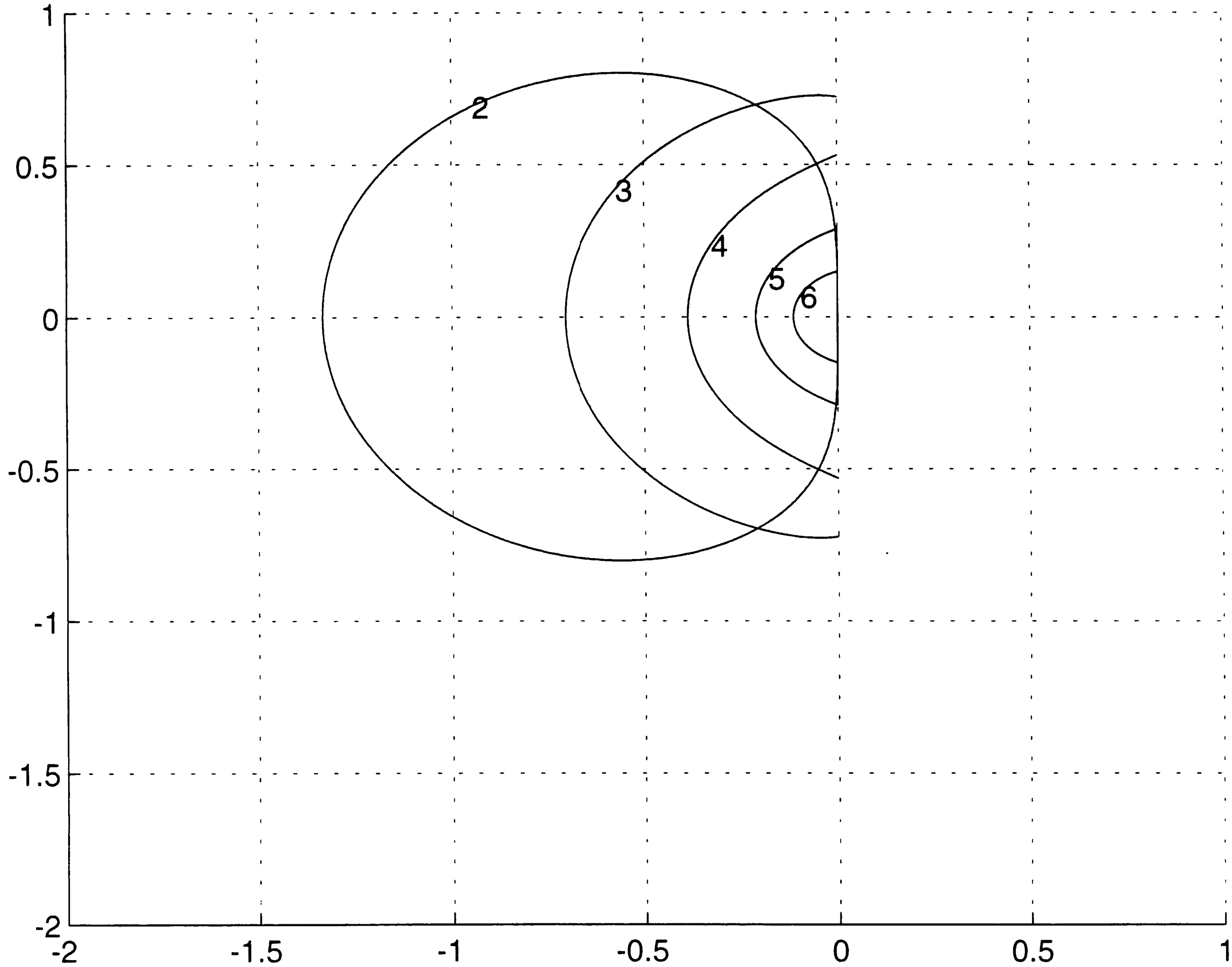

Figure 2

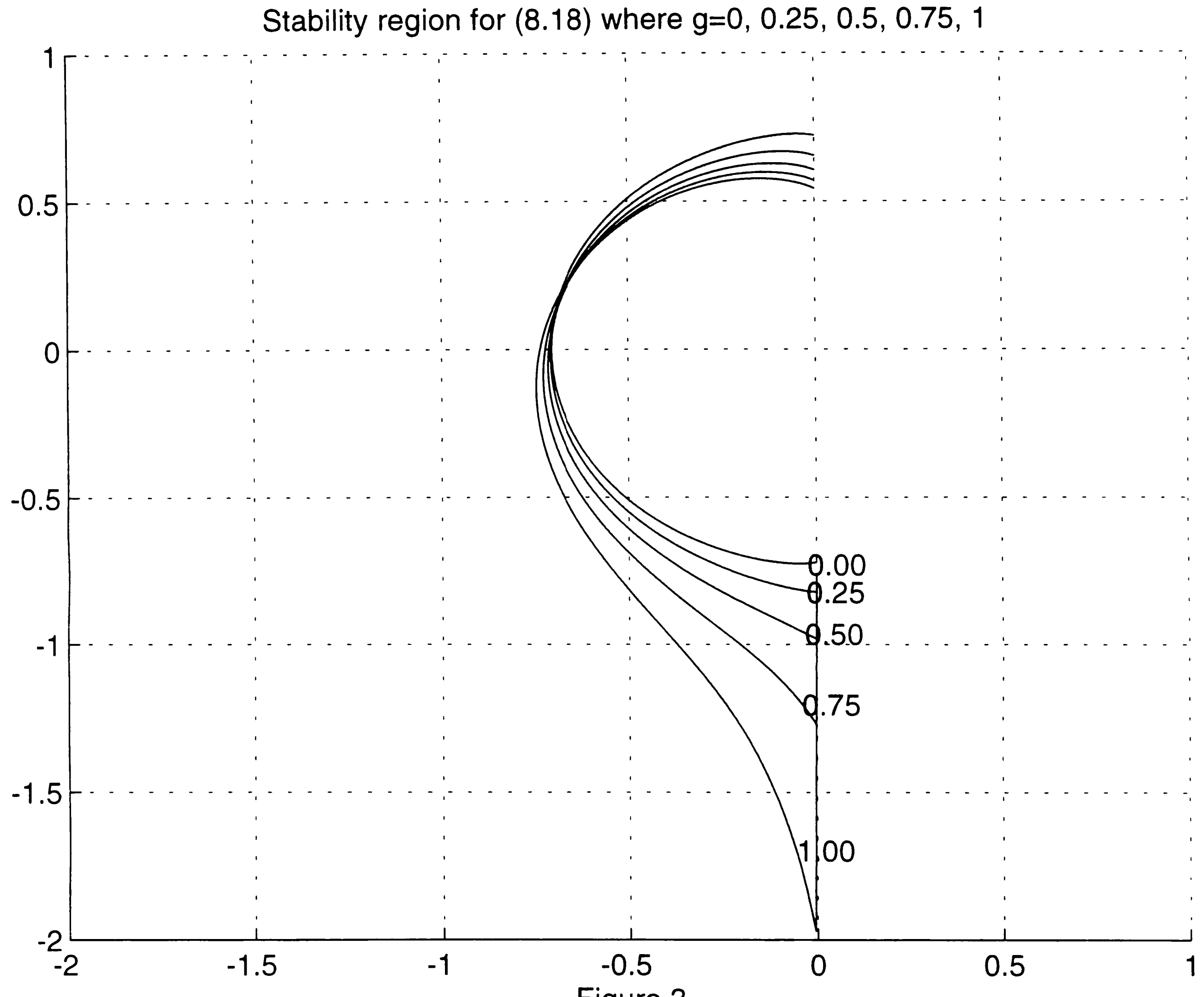
Stability region for (8.18) where g=0, 0.25, 0.5, 0.75, 1
1
0.5
0
-0.5
-1
-1.5
-2
-2
-1.5
-1
-0.5
0
0.5
1
0.00
0.25
0.50
0.75
1.00

Figure 3

Recent IMA Preprints

#	Author/s	Title

1202 **Joel Spencer**, The Erdös-Hanani conjecture via Talagrand's inequality
1203 **Zhangxin Chen**, Superconvergence results for Galerkin methods for wave propagation in various porous media
1204 **Russell Lyons, Robin Pemantle & Yuval Peres**, When does a branching process grow like its mean? Conceptual proofs of $L \log L$ criteria
1205 **Robin Pemantle**, Maximum variation of total risk
1206 **Robin Pemantle & Yuval Peres**, Galton-Watson trees with the same mean have the same polar sets
1207 **Robin Pemantle**, A shuffle that mixes sets of any fixed size much faster than it mixes the whole deck
1208 **Itai Benjamini, Robin Pemantle & Yuval Peres**, Martin capacity for Markov chains and random walks in varying dimensions
1209 **Wlodzimierz Bryc & Amir Dembo**, On large deviations of empirical measures for stationary Gaussian processes
1210 **Martin Hildebrand**, Some random processes related to affine random walks
1211 **Alexander E. Mazel & Yurii M. Suhov**, Ground states of a Boson quantum lattice model
1212 **Roger L. Fosdick & Darren E. Mason**, Single phase energy minimizers for materials with nonlocal spatial dependence
1213 **Bruce Hajek**, Load balancing in infinite networks
1214 **Petr Klouček**, The transonic flow problems stability analysis and numerical results
1215 **Petr Klouček**, On the existence of the entropic solutions for the transonic flow problem
1216 **David A. Schmidt & Chjan C. Lim**, Full sign-invertibility and symplectic matrices
1217 **Piermarco Cannarsa & Maria Elisabetta Tessitore**, Infinite dimensional Hamilton-Jacobi equations and Dirichlet boundary control problems of parabolic type
1218 **Zhangxin Chen**, Multigrid algorithms for mixed methods for second order elliptic problems
1219 **Zhangxin Chen**, Expanded mixed finite element methods for linear second order elliptic problems I
1220 **Gang Bao**, A note on the uniqueness for an inverse diffraction problem
1221 **Moshe Fridman**, A two state capital asset pricing model
1222 **Paolo Baldi**, Exact asymptotics for the probability of exit from a domain and applications to simulation
1223 **Carl Dou & Martin Hildebrand**, Enumeration and random random walks on finite groups
1224 **Jaksa Cvitanic & Ioannis Karatzas**, On portfolio optimization under "drawdown" constraints
1225 **Avner Friedman & Yong Liu**, A free boundary problem arising in magnetohydrodynamic system
1226 **Dominic Welsh**, Randomised approximation schemes for Tutte-Gröthendieck invariants
1227 **Zhangxin Chen, Bernardo Cockburn, Carl L. Gardner, & Joseph W. Jerome**, Quantum hydrodynamic simulation of hysteresis in the resonant tunneling diode
1228 **E.G. Kalnins, G.C. Williams, & Willard Miller, Jr.**, Intrinsic characterisation of the separation constant for spin one and gravitational perturbations in Kerr geometry
1229 **Zhangxin Chen**, Large-scale averaging analysis of multiphase flow in fractured reservoirs
1230 **Bruce Hajek & Babu Narayanan**, Multigraphs with the most edge covers
1231 **K.B. Athreya**, Entropy maximization
1232 **F.I. Karpelevich & Yu.M. Suhov**, Functional equations in the problem of boundedness of stochastic branching dynamics
1233 **E. Dibenedetto & V. Vespri**, On the singular equation $\beta(u)_t = \Delta u$
1234 **M.Ya. Kelbert & Yu.M. Suhov**, The Markov branching random walk and systems of reaction-diffusion (Kolmogorov-Petrovskii-Piskunov) equations
1235 **M. Hildebrand**, Random walks on random regular simple graphs
1236 **W.S. Don & A. Solomonoff**, Accuracy enhancement for higher derivatives using Chebyshev collocation and a mapping technique
1237 **D. Gurarie**, Symmetries and conservation laws of two-dimensional hydrodynamics
1238 **Z. Chen**, Finite element methods for the black oil model in petroleum reservoirs
1239 **G. Bao & A. Friedman**, Inverse problems for scattering by periodic structure
1240 **G. Bao**, Some inverse problems in partial differential equations
1241 **G. Bao**, Diffractive optics in periodic structures: The TM polarization
1242 **C.C. Lim & D.A. Schmidt**, On noneven digraphs and symplectic pairs
1243 **H.M. Soner, S.E. Shreve & J. Cvitanić**, There is no nontrivial hedging portfolio for option pricing with transaction costs
1244 **D.L. Russell & B-Yu Zhang**, Exact controllability and stabilizability of the Korteweg-de Vries equation
1245 **B. Morton, D. Enns & B-Yu Zhang**, Stability of dynamic inversion control laws applied to nonlinear aircraft pitch-axis models
1246 **S. Hansen & G. Weiss**, New results on the operator Carleson measure criterion
1247 **V.A. Malyshev & F.M. Spieksma**, Intrinsic convergence rate of countable Markov chains
1248 **G. Bao, D.C. Dobson & J.A. Cox**, Mathematical studies in rigorous grating theory

1249 **G. Bao & W.W. Symes**, On the sensitivity of solutions of hyperbolic equations to the coefficients
1250 **D.A. Huntley & S.H. Davis**, Oscillatory and cellular mode coupling in rapid directional solidification
1251 **M.J. Donahue, L. Gurvits, C. Darken & E. Sontag**, Rates of convex approximation in non-Hilbert spaces
1252 **A. Friedman & B. Hu**, A Stefan problem for multi-dimensional reaction diffusion systems
1253 **J.L. Bona & B-Y. Zhang**, The initial-value problem for the forced Korteweg-de Vries equation
1254 **A. Friedman & R. Gulliver, Organizers**, Mathematical modeling for instructors
1255 **S. Kichenassamy**, The prolongation formula for tensor fields
1256 **S. Kichenassamy**, Fuchsian equations in Sobolev spaces and blow-up
1257 **H.S. Dumas, L. Dumas, & F. Golse**, On the mean free path for a periodic array of spherical obstacles
1258 **C. Liu**, Global estimates for solutions of partial differential equations
1259 **C. Liu**, Exponentially growing solutions for inverse problems in PDE
1260 **Mary Ann Horn & I. Lasiecka**, Nonlinear boundary stabilization of parallelly connected Kirchhoff plates
1261 **B. Cockburn & H. Gau**, A posteriori error estimates for general numerical methods for scalar conservation laws
1262 **B. Cockburn & P-A. Gremaud**, A priori error estimates for numerical methods for scalar conservation laws. Part I: The general approach
1263 **R. Spigler & M. Vianello**, Convergence analysis of the semi-implicit euler method for abstract evolution equations
1264 **R. Spigler & M. Vianello**, WKB-type approximation for second-order differential equations in C^*-algebras
1265 **M. Menshikov & R.J. Williams**, Passage-time moments for continuous non-negative stochastic processes and applications
1266 **C. Mazza**, On the storage capacity of nonlinear neural networks
1267 **Z. Chen, R.E. Ewing & R. Lazarov**, Domain decomposition algorithms for mixed methods for second order elliptic problems
1268 **Z. Chen, M. Espedal & R.E. Ewing**, Finite element analysis of multiphase flow in groundwater hydrology
1269 **Z. Chen, R.E. Ewing, Y.A. Kuznetsov, R.D. Lazarov & S. Maliassov**, Multilevel preconditioners for mixed methods for second order elliptic problems
1270 **S. Kichenassamy & G.K. Srinivasan**, The structure of WTC expansions and applications
1271 **A. Zinger**, Positiveness of Wigner quasi-probability density and characterization of Gaussian distribution
1272 **V. Malkin & G. Papanicolaou**, On self-focusing of short laser pulses
1273 **J.N. Kutz & W.L. Kath**, Stability of pulses in nonlinear optical fibers using phase-sensitive amplifiers
1274 **S.K. Patch**, Recursive recovery of a family of Markov transition probabilities from boundary value data
1275 **C. Liu**, The completeness of plane waves
1276 **Z. Chen & R.E. Ewing**, Stability and convergence of a finite element method for reactive transport in ground water
1277 **Z. Chen & Do Y. Kwak**, The analysis of multigrid algorithms for nonconforming and mixed methods for second order elliptic problems
1278 **Z. Chen**, Expanded mixed finite element methods for quasilinear second order elliptic problems II
1279 **M.A. Horn & W. Littman**, Boundary control of a Schrödinger equation with nonconstant principal part
1280 **R.E. Ewing, Yu.A. Kuznetsov, R.D. Lazarov & S. Maliassov**, Substructuring preconditioning for finite element approximations of second order elliptic problems. I. Nonconforming linear elements for the Poisson equation in a parallelepiped
1281 **S. Maliassov**, Substructuring preconditioning for finite element approximations of second order elliptic problems. II. Mixed method for an elliptic operator with scalar tensor
1282 **V. Jakšić & C.-A. Pillet**, On model for quantum friction II. Fermi's golden rule and dynamics at positive temperatures
1283 **V. M. Malkin**, Kolmogorov and nonstationary spectra of optical turbulence
1284 **E.G. Kalnins, V.B. Kuznetsov & W. Miller, Jr.**, Separation of variables and the XXZ Gaudin magnet
1285 **E.G. Kalnins & W. Miller, Jr.**, A note on tensor products of q-algebra representations and orthogonal polynomials
1286 **E.G. Kalnins & W. Miller, Jr.**, q-algebra representations of the Euclidean, pseudo-Euclidean and oscillator algebras, and their tensor products
1287 **L.A. Pastur**, Spectral and probabalistic aspects of matrix models
1288 **K. Kastella**, Discrimination gain to optimize detection and classification
1289 **L.A. Peletier & W.C. Troy**, Spatial patterns described by the Extended Fisher-Kolmogorov (EFK) equation: Periodic solutions
1290 **A. Friedman & Y. Liu**, Propagation of cracks in elastic media
1291 **A. Friedman & C. Huang**, Averaged motion of charged particles in a curved strip
1292 **G. R. Sell**, Global attractors for the 3D Navier-Stokes equations
1293 **C. Liu**, A uniqueness result for a general class of inverse problems
1294 **H-O. Kreiss**, Numerical solution of problems with different time scales II
1295 **B. Cockburn, G. Gripenberg, S-O. Londen**, On convergence to entropy solutions of a single conservation law
1296 **S-H. Yu**, On stability of discrete shock profiles for conservative finite difference scheme
1297 **H. Behncke & P. Rejto**, A limiting absorption principle for separated Dirac operators with Wigner Von Neumann type potentials