Comp. & Maths. with Appls. Vol. 12A. No. 3. pp. 309–318. 1986
Printed in Great Britain.

0886–9553/86 $3.00 + .00

ON THE SHIFT PARAMETER IN THE BACKWARD BEAM METHOD FOR PARABOLIC PROBLEMS FOR PRECEDING TIMES

J. G. VERWER
Centre for Mathematics and Computer Science, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

(*Received May* 1985)

Communicated by L. F. Shampine

Abstract—We consider the backward beam method of Buzbee & Carasso [*Math. Comp.* 27, 237–267, 1973] for the numerical computation of parabolic problems for preceding times. The performance of this method is strongly influenced by the choice of a spectral shift parameter. Using logarithmic convexity arguments Buzbee & Carasso derived an expression for the optimal value for linear problems. The main concern of this paper is to illustrate that this expression can also be found and explained via the numerical stability analysis of the forward and backward recurrence involved.

1. INTRODUCTION

This paper is devoted to a study of the backward beam method of Buzbee & Carasso[2,3] for the numerical solution of parabolic problems for preceding times. Such problems are ill-posed. In [2,3] the method is discussed in the setting of an abstract Hilbert space. We consider the finite dimensional ODE system

$$\dot{U} = F(t, U), \quad 0 < t < T, \quad F(t,\cdot):\mathbf{R}^m \longrightarrow \mathbf{R}^m, \tag{1.1}$$

which is assumed to represent a continuous time, semi-discrete approximation to the parabolic PDE problem under consideration (method of lines). Given an inner product norm on $\mathbf{R}^m$, it is supposed that (1.1) is dissipative for this norm, by which we mean that the Jacobian matrix $F'(t,\cdot)$ satisfies the logarithmic norm inequality

$$\mu[F'(t, \zeta)] \leq \nu < 0, \quad \text{all } \zeta \in \mathbf{R}^m, \tag{1.2}$$

on $[0, T]$, where ν is a constant (not depending on the grid spacing used in the discretization of the space variables). For any two solutions $\tilde{U}$, U of (1.1) it then holds that ([4], [5] ch. 10; see also [7] for convergence questions)

$$\|\tilde{U}(t_2) - U(t_2)\| \leq e^{\nu(t_2-t_1)}\|\tilde{U}(t_1) - U(t_1)\|, \quad 0 \leq t_1 \leq t_2 \leq T, \tag{1.3}$$

i.e. exponential stability. Inequality (1.3) reflects the smoothing property of the parabolic equation.

Throughout our paper, $U(t)$, $0 \leq t \leq T$, represents a sufficiently smooth, exact solution of (1.1). The backward problem we examine consists of finding U on $[0, T]$ given the terminal value $U(T)$. This problem is ill-posed. Although the exact solution $U(t)$, $0 \leq t \leq T$, is a smooth function, arbitrarily small perturbations $\tilde{U}(T)$ of $U(T)$ introduce very large extraneous high frequencies. As the spatial mesh is refined, these extraneous solution components may be amplified without bound as time evolves backwards, even within arbitrarily small time intervals. To overcome the ill-posedness it is necessary to add a constraint on the set of admissible solutions. A feasible approach lies in the use of an à priori bound on the unknown initial vector $U(0)$ (see [2,3,6] for details). Following this approach our backward problem for (1.1) is recast into the initial-terminal value problem: find all solutions $\tilde{U}(t)$, $0 \leq t \leq T$, satisfying the constraints

$$\|\tilde{U}(T) - U(T)\| \leq \beta, \tag{1.4a}$$

$$\|\tilde{U}(0) - U(0)\| \leq M. \tag{1.4b}$$

The terminal bound (1.4a) accommodates error in the given data $U(T)$. M is supposed to be known from the physics of the problem. It is assumed, of course, that these conditions are compatible. If they are, the solutions U and $\tilde{U}$ can then be proven to satisfy the continuous dependence inequality

$$\|\tilde{U}(t) - U(t)\| \leq M^{1-t/T}\beta^{t/T}, \quad 0 \leq t \leq T. \tag{1.5}$$

provided the problem is linear with $F(t, U) = AU + G(t)$, A constant. This inequality is a consequence of the fact that $\|\tilde{U}(t) - U(t)\|$ is logarithmically convex[6]. One can say that due to the constraints (1.4) the unstable backward problem for (1.1) is changed into a stable initial-terminal value problem. The problem studied in [2, 3] (linear problem in abstract Hilbert space) is of this type.

In our setting the backward beam method of Buzbee and Carasso may be interpreted as a standard finite difference method for two point boundary value problems for second order ODE systems

$$\ddot{U} = H(t, U), \quad 0 < t < T, \quad H(T, U) = F_t(t, U) + F'(t, U)F(t, U). \tag{1.6}$$

The finite difference method is given by

$$U^{n+1} - 2U^n + U^{n-1} = \tau^2 H(t_n, U^n), \quad n = 1(1)N, \tag{1.7}$$

where N is a given integer, $\tau = T/(N + 1)$, and U^n is meant to approximate $U(t_n)$, the exact solution of the first order system (1.3) at time $t = t_n = n\tau$. Suppose first that the true values $U(0)$, $U(T)$ are used as boundary values, $U^0 = U(0)$ and $U^{N+1} = U(T)$. The approximation $\{U^n\}$ to the restriction $\{U(t_n)\}$ of the smooth solution U is now second order consistent and, as usual, stability then must render second order convergence. In the actual application estimates $\tilde{U}^0$ and $\tilde{U}^{N+1}$ must be implemented. Here, $\tilde{U}^{N+1}$ is considered as a perturbation of $U(T)$ as in the formulation above, while $\tilde{U}^0$ stands for a cruder estimate of the unknown initial vector $U(0)$. In what follows the estimates $\tilde{U}^{N+1}$ and $\tilde{U}^0$ are supposed to satisfy (1.4). In application we thus have time integration errors $\tilde{U}^n - U(t^n)$ consisting of two parts, one part being due to truncation and the other caused by using the wrong boundary values. The motivation for this approach is that the errors due to using the wrong boundary values are damped when going into the interior of $[0, T]$.

Theoretically these numerical boundary errors should obey stability inequalities similar to (1.5) for the continuous time backward problem. For interesting classes of linear problems Buzbee and Carasso did, in fact, succeed in deriving error bounds which differ from the fundamental uncertainty (1.5) by only the contribution due to the truncation. Among others, for the linear problem

$$\dot{U} = AU + G(t), \quad \text{A symmetric negative definite,} \tag{1.8}$$

where A is assumed to be independent of t, they recovered the inequality (1.5) for the numerical boundary errors (see [2], p. 253 or [3], p. 132). Their analysis is based on the spectral shift transformation

$$V(t) = e^{kt}U(t), \quad k \in \mathbf{R}, \tag{1.9}$$

which transforms (1.8) into

$$\dot{V} = (A + kI)V + e^{kt}G(t). \tag{1.10}$$

An analytical stability analysis of the corresponding second order form (the backward beam equation) then leads to the numerical continuous dependence inequality if the shift parameter k is chosen as

$$k = \frac{1}{T}\ln\frac{M}{\beta}. \tag{1.11}$$

The purpose of this paper is to show that this optimal expression for k, optimal in the sense that it yields the maximal overall damping of the boundary errors, can also be found via a numerical stability analysis of the forward and backward recurrence involved. Our alternative derivation of (1.11) is longer than that in [2,3]. However, the stability analysis of the forward and backward recurrence provides more insight into the numerical process and the role of the transformation (1.9). Among other things, it shows that the shift provides no real practical advantage when k is not larger than minus the spectral abscissa of A. We also derive a dependence inequality like (1.5) for the discrete variables. As contrasted to that of [2], where τ is supposed to be sufficiently small, our derivation is valid for any $\tau > 0$.

2. THE BACKWARD BEAM ANALYSIS

By way of comparison we shall first sketch the so-called backward beam derivation as presented in [2,3]. Let $\tilde{U}^0$, $\tilde{U}^{N+1}$ be the estimates of values $U(0)$, $U(T)$ which lie on an exact solution $U(t)$, $0 \leq t \leq T$, of the linear problem (1.8). Let $\tilde{U}(t) = \exp(-kt)\tilde{V}(t)$, $0 \leq t \leq T$, where $\tilde{V}$ is the exact solution of the two point boundary value problem

$$\ddot{\tilde{V}} = (A + kI)^2\tilde{V} + \text{inh. term}, \quad \tilde{V}(0) = \tilde{U}^0, \quad \tilde{V}(T) = e^{kT}\tilde{U}^{N+1}. \tag{2.1}$$

Likewise, we consider the two point problem for $V(t) = \exp(kt)U(t)$, i.e.,

$$\ddot{V} = (A + kI)^2V + \text{inh. term}, \quad V(0) = U(0), \quad V(T) = e^{kT}U(T). \tag{2.2}$$

Let $\|\cdot\|_2$ be the Euclidean norm and $\langle\cdot,\cdot\rangle_2$ the standard inner product. Then $W = \tilde{V} - V$ satisfies

$$\frac{d^2}{dt^2}\|W(t)\|_2^2 = 2\|\dot{W}(t)\|_2^2 + 2\langle(A + kI)^2W(t), W(t)\rangle_2 \geq 0, \tag{2.3}$$

i.e. W is norm-convex. The convexity implies that

$$\|W(t)\|_2 \leq \frac{T - t}{T}\|W(0)\|_2 + \frac{t}{T}\|W(T)\|_2, \quad 0 \leq t \leq T, \tag{2.4}$$

or, equivalently,

$$\|\tilde{U}(t) - U(t)\|_2 \leq e^{-kt}\frac{T - t}{T}M + e^{k(T-t)}\frac{t}{T}\beta, \quad 0 \leq t \leq T, \tag{2.5}$$

provided $\tilde{U}^0$ and $\tilde{U}^{N+1}$ satisfy the constraints (1.4). Substitution of the expression (1.11) for k yields

$$\|\tilde{U}(t) - U(t)\|_2 \leq M^{1-t/T}\beta^{t/T}, \quad 0 \leq t \leq T, \tag{2.6}$$

i.e. a numerical continuous dependence inequality similar to (1.5). It is emphasized that here $\tilde{U}$ is related to the exact solution $\tilde{V}$ of (2.1).

In conclusion, if we are able to solve the two point problem (2.1) sufficiently accurately, we end up with a numerical solution $\tilde{U}^n$ to the initial-terminal value problem for (1.8) which satisfies the stability inequality (2.6), except for truncation errors. Throughout it is assumed that these latter errors can be made sufficiently small. Here we should mention that the spectral shift transformation usually leads to larger truncation errors. In practice this means that the transformation may force us to use smaller values of the time step τ. Buzbee and Carasso used the difference scheme (1.7) for the numerical solution of (2.1). Other numerical techniques may also be considered.

3. THE FORWARD AND BACKWARD RECURRENCE

Let $\tilde{V}^n$, V^n, $n = 1(1)N$, be the numerical approximations, defined by the finite difference method (1.7), to the exact solutions $\tilde{V}(t_n)$, $V(t_n)$ of the two point problems (2.1), (2.2). Let $W^n = \tilde{V}^n - V^n$; W^n satisfies

$$\begin{aligned} W^{n+1} - 2W^n + W^{n-1} &= \tau^2(A + kI)^2W^n, \quad n = 1(1)N, \\ W^0 = \tilde{U}^0 - U(0), &\quad W^{N+1} = e^{kT}(\tilde{U}^{N+1} - U(T)). \end{aligned} \tag{3.1}$$

The total numerical error to be examined is given by

$$e^{-kt_n}\tilde{V}^n - U(t_n) = e^{-kt_n}W^n e^{-kt_n}(V^n - U(t_n)), \tag{3.2}$$

the second part of which is due to truncation and the first part due to using the wrong boundary values $\tilde{U}^0$, $\tilde{U}^{N+1}$. In what follows we shall examine the propagation of the boundary errors $e^{-kt_n}W^n$ by studying the stability of the forward and backward recurrences which arise in the solution of (3.1).

Firstly, (3.1) is rewritten in the block tridiagonal matrix form

$$(E \otimes I - \tau^2 \operatorname{diag}(A + kI)^2)\vec{W} = \vec{R}, \tag{3.3}$$

$$E = \begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix}_{N \times N}, \quad \vec{W} = \begin{bmatrix} W^N \\ W^{N-1} \\ \vdots \\ W^2 \\ W^1 \end{bmatrix},$$

$$\vec{R} = \begin{bmatrix} -e^{kT}(\tilde{U}^{N+1} - U(T)) \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ -(\tilde{U}^0 - U(0)) \end{bmatrix},$$

where we have reversed the order of $W^1, W^2, \ldots, W^N$. Secondly, we decompose the matrix in (3.3) as LU, viz.,

$$L = \begin{bmatrix} I & & & \\ D_1^{-1} & I & & \\ & \ddots & \ddots & \\ & & D_{N-2}^{-1} & I \\ & & & D_{N-1}^{-1} \quad I \end{bmatrix}, \quad U = \begin{bmatrix} D_1 & I & & \\ & D_2 & I & \\ & & \ddots & \ddots \\ & & & D_{N-1} \quad I \\ & & & \quad D_N \end{bmatrix}, \tag{3.4}$$

where the $m \times m$ matrices D_n are given by

$$\begin{aligned} D_1 &= -2I - \tau^2(A + kI)^2, \\ D_{n+1} &= -2I - \tau^2(A + kI)^2 - D_n^{-1}, \quad n = 1(1)N - 1. \end{aligned} \tag{3.5}$$

Next we write (3.3) as $L\vec{Y} = \vec{R}$, $U\vec{W} = \vec{Y}$ where $\vec{Y} = [(Y^N)^T, \ldots, (Y^1)^T]^T$, so that we arrive at the backward recurrence (decreasing n)

$$\begin{aligned} Y^N &= -e^{kT}(\tilde{U}^{N+1} - U(T)), \\ Y^n &= -D_{N-n}^{-1}Y^{n+1}, \quad n = N - 1(-1)2, \\ Y^1 &= -D_{N-1}^{-1}Y^2 - (\tilde{U}^0 - U(0)), \end{aligned} \tag{3.6}$$

followed by the forward recurrence (increasing n)

$$\begin{aligned} W^1 &= D_N^{-1}Y^1, \\ W^n &= -D_{N-n+1}^{-1}W^{n-1} + D_{N-n+1}^{-1}Y^n, \quad n = 2(1)N. \end{aligned} \tag{3.7}$$

These two recurrences describe the propagation of the intermediate boundary errors W^n into the interior of $[0, T]$. There remains the back transformation $\exp(-kt_n)W^n$. Let $\hat{W}^n = \exp(-kt_n)W^n$ and $\hat{Y}^n = \exp(-kt_n)Y^n$. Then the final backward recurrence reads

$$\begin{aligned} \hat{Y}^N &= -(\bar{U}^{N+1} - U(T)), \\ \hat{Y}^n &= -(e^{k\tau}D_{N-n}^{-1})\hat{Y}^{n+1}, \quad n = N - 1(-1)2, \\ \hat{Y}^1 &= (e^{k\tau}D_{N-1}^{-1})\hat{Y}^2 - e^{-k\tau}(\bar{U}^0 - U(0)), \end{aligned} \tag{3.8}$$

and the final forward recurrence is given by

$$\begin{aligned} \hat{W}^1 &= D_N^{-1}\hat{Y}^1, \\ \hat{W}^n &= -(D_{N-n+1}^{-1}e^{-k\tau})\hat{W}^{n-1} + D_{N-n+1}^{-1}\hat{Y}^n, \quad n = 2(1)N. \end{aligned} \tag{3.9}$$

We see that the propagation of the terminal error $\bar{U}^{N+1} - U(T)$ takes place in the backward and forward recurrences, whilst the initial error $\bar{U}^0 - U(0)$ is propagated only in the forward direction. Observe that the recurrence (3.5) for the amplification matrices is well-defined, as A is symmetric negative definite.

4. THE DISCRETE DEPENDENCE INEQUALITY

We shall derive a dependence inequality for the internal boundary errors $\hat{W}^n$ which is analogous to (2.6). The interesting thing about this derivation, in comparison to that of section 2 and of [2], p. 253, is that the dependence inequality turns out to be valid for any value of $\tau > 0$.

Consider the backward recurrence (3.8). The following inequalities hold

$$\begin{aligned} &\|\hat{Y}^N\|_2 \leqslant \beta, \\ &\|\hat{Y}^n\|_2 \leqslant e^{k\tau}\|D_{N-n}^{-1}\|_2\|\hat{Y}^{n+1}\|_2, \quad n = N - 1(-1)2, \\ &\|\hat{Y}^1\|_2 \leqslant e^{k\tau}\|D_{N-1}^{-1}\|_2\|\hat{Y}^2\|_2 + e^{-k\tau}M. \end{aligned} \tag{4.1}$$

As A is symmetric, each matrix D_n^{-1} is symmetric so that its spectral norm is equal to its spectral radius. It thus follows that

$$\|D_n^{-1}\|_2 \leqslant \frac{n}{n+1}, \quad n = 1(1)N, \tag{4.2}$$

with equality if $-k$ is an eigenvalue of A (the eigenvalues of D_n satisfy recurrence (3.5), too). Substitution of (4.2) into (4.1) yields

$$\begin{aligned} \|\hat{Y}^n\|_2 &\leqslant e^{k\tau(N-n)}\frac{\beta}{N-n+1}, \quad n = N(-1)2, \\ \|\hat{Y}^1\|_2 &\leqslant e^{k\tau(N-1)}\frac{\beta}{N} + e^{-k\tau}M, \end{aligned} \tag{4.3}$$

for the intermediate variables $\hat{Y}^n$.

Next consider the forward recurrence (3.9). Substitution of (4.2), (4.3) gives

$$\begin{aligned} \|\hat{W}^1\|_2 &\leqslant \frac{N}{N+1}\|\hat{Y}^1\|_2 \leqslant e^{k\tau(N-1)}\frac{\beta}{N+1} + e^{-k\tau}\frac{N}{N+1}M, \\ \|\hat{W}^n\|_2 &\leqslant \frac{N-n+1}{N-n+2}e^{-k\tau}\|\hat{W}^{n-1}\|_2 + e^{k\tau(N-n)}\frac{\beta}{N-n+2}, \quad n = 2(1)N. \end{aligned} \tag{4.4}$$

An elementary computation then results in

$$\begin{aligned} \|\hat{W}^n\|_2 &\leqslant e^{-nk\tau}\frac{N-n+1}{N+1}M + e^{(N-n)k\tau}\frac{n}{N+1}\beta, \\ &\leqslant e^{-nk\tau}\frac{N-n+1}{N+1}M + e^{(N-n+1)k\tau}\frac{n}{N+1}\beta, \quad n = 0(1)N+1, \end{aligned} \tag{4.5}$$

if $k \geqslant 0$. Substitution of $t_n = n\tau$, where $\tau = T/(N+1)$, yields

$$\|\hat{W}^n\|_2 \leqslant e^{-kt_n}\frac{T-t_n}{T}M + e^{k(T-t_n)}\frac{t_n}{T}\beta, \quad n = 0(1)N+1. \tag{4.6}$$

This formula is equivalent to (2.5), so that, after substitution of $k = T^{-1}\ln(M/\beta) > 0$, we arrive at the discrete dependence inequality

$$\|\hat{W}^n\|_2 \leqslant M^{1-t_n/T}\beta^{t_n/T}, \quad n = 0(1)N+1. \tag{4.7}$$

The inequalities (4.6) and (2.5) indicate that with a positive shift the forward damping may improve, but also that the terminal error may be amplified. Because β is supposed to be smaller than M, some backward amplification is allowed. This leads us to the problem of determining the optimal value of k, which is simply the maximal value under the constraint just mentioned. As it turns out, $k = T^{-1}\ln(M/\beta)$ is optimal in the sense that for this value the discrete dependence on data inequality (4.7) holds for any $\tau > 0$.

This result does not necessarily imply that for this particular value of k the true overall damping of boundary errors is always improved by the transformation (1.9). In the remainder of this paper we shall try to provide insight into this matter by carrying out a precise spectral analysis of the recurrences (3.8), (3.9). Among other things, we show that for problem (1.8) the transformation provides no real advantage when $k \leqslant -\alpha[A]$, $\alpha[A]$ being the spectral abscissa of A.

5. THE PROPAGATION OF THE BOUNDARY ERRORS

The symmetric negative definiteness of A implies that A is othogonally similar to its eigenvalue matrix $\text{diag}(\delta_j) = X^{-1}AX$ with all $\delta_j < 0$. This is also true for all the matrices D_n. Hence we may work with the eigencomponent recurrences (backward)

$$\begin{aligned} \hat{y}_N &= -\epsilon_{N+1}, \\ \hat{y}_n &= -(e^{k\tau}d_{N-n}^{-1})\hat{y}_{n+1}, \quad n = N-1(-1)2, \\ \hat{y}_1 &= -(e^{k\tau}d_{N-1}^{-1})\hat{y}_2 - e^{-k\tau}\epsilon_0, \end{aligned} \tag{5.1}$$

and (forward)

$$\begin{aligned} \hat{w}_1 &= d_N^{-1}\hat{y}_1, \\ \hat{w}_n &= -(d_{N-n+1}^{-1}e^{-k\tau})\hat{w}_{n-1} + d_{N-n+1}^{-1}\hat{y}_n, \quad n = 2(1)N. \end{aligned} \tag{5.2}$$

To distinguish from the vector case we use here subscripts instead of superscripts. The eigenvalues d_n of D_n satisfy

$$d_1 = -2 - \tau^2(\delta + k)^2, \quad d_{n+1} = -2 - \tau^2(\delta + k)^2 - d_n^{-1}, \quad n = 1(1)N - 1, \tag{5.3}$$

δ being generic for δ_j, $j = 1(1)m$ (note that the eigencomponent $\hat{w}_n$ is the j-th component of $X^{-1}\hat{W}^n$ if $\delta = \delta_j$). The propagation of the boundary errors, now represented by the eigencomponents ϵ_0 and ϵ_{N+1}, is fully described by (5.1)–(5.3). The eigenvalues d_n determine the damping or amplification of ϵ_0 and ϵ_{N+1}.

Consider recurrence (5.3). Each d_n is a continued fraction in $-2 - \tau^2(\delta + k)^2$ and thus converges (see, e.g., [1], p. 19) to a limit d, as $n \to \infty$, which satisfies the quadratic equation $d^2 + (2 + \tau^2(\delta + k)^2)d + 1 = 0$. As $\tau^2(\delta + k)^2 \geq 0$, we find

$$d = -1 - \frac{1}{2}\tau^2(\delta + k)^2 - \frac{1}{2}\sqrt{(2 + \tau^2(\delta + k)^2)^2 - 4}. \tag{5.4}$$

The convergence of d_n to d is monotone and $0 < -d_n^{-1} < -d^{-1} \leq 1$ for all $n = 1(1)N$ and all $\tau^2(\delta + k)^2$. This means that the amplification factors $-e^{k\tau}d_{N-n}^{-1}$ in (5.1) and $-e^{-k\tau}d_{N-n+1}^{-1}$ in (5.2) are majorized by $-e^{k\tau}d^{-1}$ and $-e^{-k\tau}d^{-1}$, respectively.

It is emphasized that the rate of convergence of d_n to d depends on the size of $\tau^2(\delta + k)^2$. The larger this number, the faster the convergence, so that it is slowest for $\delta = -k$. A consequence is that for values of δ close to $-k$ the insertion of the upper bound $-d^{-1}$ for $-d_n^{-1}$ into the recurrences will lead to somewhat pessimistic conclusions. It is also of interest to note that $-d^{-1}$ monotonically decreases as $\tau^2(\delta + k)^2$ increases. This also holds for $-d_n^{-1}$ for all n. This property implies trivially that for $k = 0$ high frequent error components are damped faster than low frequency ones (in both directions; note that $-d^{-1} < 1$ for $\tau^2\delta^2 > 0$). Hence for $k = 0$ the spectral abscissa of A will be decisive for the damping of the boundary errors in (3.6) and (3.7).

The limit value d satisfies the asymptotic relation

$$-d^{-1} = e^{-\tau|\delta + k|} + O(\tau^3|\delta + k|^3), \quad \tau \longrightarrow 0. \tag{5.5}$$

This leads us to the definitions of the asymptotic forward amplification factor,

$$FAF = e^{\tau(-k-|\delta+k|)}, \tag{5.6}$$

and the asymptotic backward amplification factor,

$$BAF = e^{\tau(k-|\delta+k|)}. \tag{5.7}$$

For values of δ not too close to $-k$, *FAF* and *BAF* are accurate substitutes for the true amplification factors in almost all stages $n = 1, 2, \cdots$. For n close to N (near the terminal point T) these factors may be a bit crude due to the fact that for these stages d_{N-n} is still too far away from d. However, recall that $-d^{-1} \geq -d_n^{-1}$ for all n. This is true for all $\tau^2(\delta + k)^2 \geq 0$, so that if $\delta \simeq -k$, the asymptotic factor will overestimate the true factors. Consequently, in all cases *FAF* and *BAF* are safe substitutes, provided τ is sufficiently small. These factors are useful for illustrating the effect of the transformation (1.9) on the boundary error propagation.

6. THE EFFECT OF THE SPECTRAL SHIFT TRANSFORMATION

In this section we examine *FAF* and *BAF* for $\delta + k > 0$ and $\delta + k < 0$. As we mentioned, for δ close to $-k$ the true factors may be smaller, implying that in this range the conclusions are a bit loose.

We first examine FAF for which

$$FAF = \begin{cases} e^{-\tau\delta - 2\tau k}, & \delta + k > 0, \\ e^{\tau\delta}, & \delta + k \leq 0. \end{cases} \tag{6.1}$$

It follows that for $\delta + k < 0$, the shift has no influence on the forward damping. However, there is a change if $\delta + k > 0$. Then

$$e^{\tau(-\delta - 2k)} = e^{\tau\delta} e^{\tau(-2\delta - 2k)} < e^{\tau\delta} \quad \text{if} \quad \delta + k > 0. \tag{6.2}$$

which implies that the damping is accelerated for all δ. k satisfying $\delta + k > 0$. The acceleration factor is $\exp(-2\tau(\delta + k))$. Concerning the forward damping, for problem (1.8) the shift provides a real advantage only if $k > -\alpha[A] > 0$, and the larger k, the greater the acceleration.

In the backward direction the situation is less favourable. We have

$$BAF = \begin{cases} e^{-\tau\delta}, & \delta + k > 0, \\ e^{\tau(\delta + 2k)}, & \delta + k \leq 0. \end{cases} \tag{6.3}$$

Hence for problem (1.8) the shift influences the backward damping of all error components ϵ_{N+1}. Those components for which $\delta + k > 0$ are now amplified by the factor $e^{-\tau\delta}$ (without shift they are damped by the factor $e^{\tau\delta}$). The components for which $\delta + k < 0$ are propagated with the factor $e^{\tau(\delta + 2k)}$. So, for $k < 0$ we get an improved backward damping. Unfortunately, for $k < 0$ the forward damping remains unchanged. To sum up, for $k > 0$ part of the spectrum ($\delta > -2k$) will suffer from backward amplification. Figure 1 illustrates the situation for $k > -\alpha[A]$.

We see that for a given k the backward amplification is maximal for that eigencomponent which corresponds to the eigenvalue closest to $-k$. Supposing that $\delta = -k$, $BAF = e^{\tau k}$ and the amplified backward error at $t = 0$, using BAF, is

$$e^{\tau k(N+1)} |\epsilon_{N+1}| = e^{kT} |\epsilon_{N+1}|. \tag{6.4}$$

If we require that this error be less than or equal to $|\epsilon_0|$, we find the condition

$$K \leq \frac{1}{T} \ln \frac{|\epsilon_0|}{|\epsilon_{N+1}|}, \tag{6.5}$$

which is similar to (1.11).

Let us summarize now the results of sections 5 and 6. Concerning the forward damping, which is crucial to the damping of the initial error $\tilde{U}^0 - U(0)$, the shift provides a real advantage only when $-k < \alpha[A]$. The maximal value of FAF on the spectrum of A is

$$FAF_{\max} = \begin{cases} e^{\tau(-\alpha[A] - 2k)}, & -k < \alpha[A], \\ e^{\tau\alpha[A]}, & -k \geq \alpha[A]. \end{cases} \tag{6.6}$$

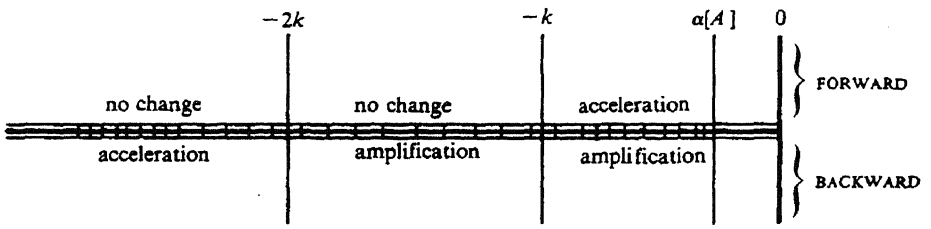

Fig. 1.

For τ sufficiently small and k not too close to $\alpha[A]$, this factor determines the decay of the initial error $\tilde{U}^0 - U^0$ at all points t_n, except for a few close to $t = T$. The maximal value of *BAF* on the spectrum of A is

$$BAF_{\max} = \begin{cases} e^{\tau k}, & -k < \alpha[A], \\ e^{\tau(\alpha[A]+2k)}, & -k \geq \alpha[A]. \end{cases} \tag{6.7}$$

This says that the shift causes backward amplification unless $k < -(1/2)\alpha[A]$. However, in applications this range of k is of no use. The shift parameter should lie approximately in the interval

$$-\alpha[A] < k \leq \frac{1}{T} \ln \frac{M}{\beta} \tag{6.8}$$

in order to improve noticeably upon the overall damping of the boundary errors without shifting. The lower bound for k is necessary to accelerate the forward damping, while the upper bound prevents the terminal error from growing too fast in the backward recurrence. Although these amplified errors are damped again in the forward direction, there is no advantage in choosing a too large value of k. To conclude, the optimal value for k is $T^{-1} \ln (M/\beta)$. However, if $T^{-1} \ln (M/\beta) < -\alpha[A]$, not much will be gained in comparison to the choice $k = 0$ (no shift).

For purposes of illustration we solved the backward problem for the simple scalar equation

$$\dot{U} = \delta(U - 1), \quad 0 \leq t \leq T, \quad U(1) = 1, \tag{6.9}$$

with exact solution $U(t) = 1$, $0 \leq t \leq T$. We used the estimates $\tilde{U}^0 = .9$, $\tilde{U}^{N+1} = .999$ so that $k_{opt} = \ln (100) \simeq 4.6$. For the range of shift values $k = 0(1)8$, Table 1 shows some results for $\delta = -1$ and -5, respectively. Here we used the very small stepsize $\tau = .005$ to prevent possible interference of truncation errors. Note that for $\delta = -5$ nothing is gained in comparison to $k = 0(k_{opt} < 5)$. This is in full agreement with our asymptotic spectral analysis. For $\delta = -1$ the relative gain in accuracy is clearly seen. However, for $k > k_{opt}$ the boundary errors do not increase, rather remain more or less on the same level. This indicates that on a large part of the interval *BAF* is too pessimistic.

In connection with the foregoing the following remark is of interest. Suppose that for the nonlinear problem (1.1) the parameters ν, T, M and β satisfy

$$T^{-1} \ln \frac{M}{\beta} < -\nu. \tag{6.10}$$

Table 1. Results for problem (6.7). The entries in the table represent $-\log_{10}$ (absolute error)

					$\delta = -1$				
$t \backslash k$	0	1	2	3	4	5	6	7	8
.2	1.12	1.18	1.29	1.43	1.58	1.74	1.89	2.04	2.19
.4	1.26	1.39	1.59	1.85	2.12	2.36	2.55	2.69	2.80
.6	1.45	1.64	1.93	2.27	2.56	2.73	2.82	2.87	2.91
.8	1.75	2.00	2.34	2.65	2.84	2.90	2.93	2.94	2.96
					$\delta = -5$				
$t \backslash k$	0	1	2	3	4	5	6	7	8
.2	1.43	1.43	1.43	1.43	1.42	1.40	1.38	1.36	1.36
.4	1.87	1.87	1.86	1.85	1.83	1.80	1.77	1.77	1.77
.6	2.30	2.29	2.28	2.27	2.24	2.20	2.18	2.18	2.20
.8	2.71	2.70	2.68	2.65	2.63	2.61	2.60	2.60	2.61

Then, as a consequence of (1.3), the continuous dependence inequality (1.5) is automatically satisfied for any pair of solutions satisfying only (1.4b). This is in line with our observation that if $k_{opt} > -\alpha[A]$ the spectral shift does not have much effect. If (6.10) holds, the initial-terminal problem can in fact be solved by any accurate, stable forward in time integration starting from $\tilde{U}^0$.

The above has led us to the conjecture that when the backward beam method is applied without shift, the numerical backward beam solution itself very often can be approximated accurately by any stable forward in time integration starting from $\tilde{U}^0$. We verified this with success on three backward heat problems, one linear and two nonlinear. The explanation lies in the fact that in such cases the initial error decay is determined mainly by the stability of the problem, a property which the backward beam method shares with any stable integration formula[5].

Acknowledgements—The numerical experiments were carried out by Mrs. M. Louter-Nool who is gratefully acknowledged for her assistance. It is also a pleasure to acknowledge several stimulating discussions with Owe Axelsson.

REFERENCES

1. M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover. New York (1972).
2. B. L. Buzbee, and A. Carasso. On the numerical computation of parabolic problems for preceding times. *Math. Comp.* **27**, 237–266. (1973).
3. A. Carasso, The backward beam equation and the numerical computation of dissipative equations backwards in time, in *Improperly posed boundary value problems*. (A. Carasso & A. P. Stone eds.). Research Notes in Math. 1. Pitman, London-San Francisco-Melbourne (1975).
4. G. Dahlquist, Stability and error bounds in the numerical integration of ordinary differential equations. *Trans. Royal Inst. of Technology*, No 130, Stockholm (1959).
5. K. Dekker and J. G. Verwer, Stability of Runge-Kutta methods for stiff nonlinear differential equations. North-Holland Publishing Co. (1984).
6. L. E. Payne, Improperly posed problems in partial differential equations. CBMS Regional Conference Series No. 22, SIAM Publications (1975).
7. J. G. Verwer and J. M. Sanz- Serna, Convergence of method of lines approximations to partial differential equations. *Computing* **33**, 297–313 (1984).