MATHEMATICS OF COMPUTATION
VOLUME 39, NUMBER 159
JULY 1982, PAGES 69-84

Time Discretization in the Backward Solution of Parabolic Equations. II*

By Lars Eldén

Abstract. The backward beam method for solving a parabolic partial differential equation backward in time is studied.

Time discretizations based on Padé approximations of the exponential function are considered, and a priori estimates of the step length are given, which guarantee an almost optimal error bound. The computational efficiency of different discretizations is discussed. Some numerical examples are given, which compare the backward beam method and the regularization method studied in Part I of this paper.

1. Introduction. The problem of solving a parabolic partial differential equation backward in time is a classical ill-posed problem; the solution (if it exists) does not depend continuously on the data.

We write the equation in the following form

$$\begin{cases} u_t = -Lu, & 0 \leqslant t \leqslant 1, \\ u(1) = w, \end{cases} \tag{1.1}$$

where $w(x)$ is a given function in $L^2(\Omega)$, and Ω is a bounded domain in R^n with a smooth boundary $\partial\Omega$. L is the unbounded, nonnegative operator in $L^2(\Omega)$ corresponding to a selfadjoint, elliptic boundary value problem in Ω with zero Dirichlet data on $\partial\Omega$. The coefficients of L are assumed to be smooth and independent of time.

Continuous dependence on the data is restored, if we impose a bound on the solution at $t = 0$ and allow for some imprecision in the data. Thus consider the following constrained problem.

Find any solution

$$\begin{cases} u_t = -Lu, & 0 \leqslant t \leqslant 1, \\ \|u(1) - w\| \leqslant \delta, \\ \|u(0)\| \leqslant M, \end{cases} \tag{1.2}$$

where the norm is the $L^2(\Omega)$-norm, and δ and M are given positive constants, $\delta \ll M$. (Throughout this paper we assume that δ and M have been chosen so that there exist solutions of (1.2).) Using logarithmic convexity [1], [8, p. 11], it is easy to show that any two solutions of (1.2), u_1 and u_2, satisfy

$$\|u_1(t) - u_2(t)\| \leqslant 2\delta^t M^{1-t}. \tag{1.3}$$

Thus, for $0 < t \leqslant 1$, we have continuous dependence on the data.

Received April 30, 1981; revised October 8, 1981.
1980 *Mathematics Subject Classification*. Primary 65M30.
*This work was supported by the Swedish Natural Science Research Council.

Now if we want to solve (1.2), the best we can hope for, in view of (1.3), is to find an approximate solution $v(t)$ with the error estimate

$$\|u(t) - v(t)\| \leq C\delta^t M^{1-t}, \tag{1.4}$$

where $u(t)$ denotes an arbitrary solution of (1.2) and C is some constant. There are at least two methods for approximating solutions of (1.2) which have the estimate (1.4) with $C = 1$. One is the *regularization method with time-dependent regularization parameter* [9], [7], [5]. The second method is the *backward beam method* [2].

In [5] we investigated how to discretize in the regularization method, so that we get the same type of error estimate (1.4) for the time-discrete version of the method. In this paper we consider the same problem for the backward beam method. We show that it is possible to make a time discretization in this method, so that for the approximate solution $v_a(t)$ we get the error estimate

$$\|u(t) - v_a(t)\| \leq 2\delta^t M^{1-t}, \tag{1.5}$$

with u as in (1.4). We also give a priori estimates of the step length in the time discretization, which guarantee that (1.5) holds. These results will have significance for the practical solution of problems in two or more space dimensions, where the geometry is nonrectangular or the coefficients of the differential operator are nonconstant, since for such problems it is necessary to discretize in time and space. To a great extent the results of this paper parallel those in [5] for the regularization method.

The backward beam method was introduced in [2] for a more general class of equations. In Section 2 we give a brief account of the method applied to the problem (1.1)

Time discretization is discussed in rather general terms in Section 3, and preliminary error estimates are given. In Section 4 we describe more specifically a scheme based on Padé approximations of the exponential function. This time discretization is conceptually different from that suggested in [2], and its main advantage is that we can derive a priori estimates of the step length, for which (1.5) holds. Also, it requires less storage.

In Section 5 we discuss the numerical realization of our procedure, and we show that we have to solve a sequence of equations of the type

$$\left(\alpha_i L^2 + \beta_i L + \gamma_i I\right) v_i = w_i. \tag{1.6}$$

The number of equations (1.6) is different for different Padé approximations and can be taken as a measure of the efficiency of a certain approximation.

Some numerical examples for the backward beam and regularization methods are given in Section 6.

Unless otherwise stated, the norm $\|\cdot\|$ is the $L^2(\Omega)$-norm. Throughout we shall write $\exp(-Lt)$ to denote a member of the strongly continuous semigroup generated by L; see, e.g., [6]. This semigroup is easily defined in terms of the spectral representation of L.

2. The Backward Beam Method. The backward beam method for approximating solutions of parabolic equations backwards in time was introduced by Buzbee and Carasso [2]. They considered the problem where the coefficients of the elliptic operator may depend on time. In this section we give a brief description of the method, where we take advantage of our having restricted ourselves to the time-independent case.

In this section we shall consider (1.1) as an abstract ordinary differential equation (with independent variable t) on a Hilbert space. This is fully justified by our assumptions on the operator L, since these assumptions guarantee the existence of a complete set of eigenfunctions of L, and by expansion in these eigenfunctions we can replace (1.1) by an infinite system of ordinary differential equations.

Let $u(t)$ be any solution of $u_t = -Lu$, and put

$$z(t) = e^{\kappa t}u(t), \qquad \kappa = \log(M/\delta). \tag{2.1}$$

Then it is easily seen that z satisfies

$$z_t = (-L + \kappa I)z,$$

where I is the identity operator. Differentiating with respect to t, we get the second order equation

$$z_{tt} = (-L + \kappa I)^2 z. \tag{2.2}$$

Next consider the following boundary value problem for (2.2)

$$\begin{cases} z_{tt} = (-L + \kappa I)^2 z, \\ z(0) = 0, \\ z(1) = e^{\kappa}w. \end{cases} \tag{2.3}$$

This is a well-posed problem with the solution

$$z(t) = e^{\kappa}\frac{\sinh((-L + \kappa I)t)}{\sinh(-L + \kappa I)}w \tag{2.4}$$

(for notational convenience we allow ourselves to write the operator

$$(\sinh(-L + \kappa I))^{-1}(\sinh((-L + \kappa I)t))$$

as in (2.4). Note that the operator is easily defined in terms of the spectral representation of L).

To make up for the change of variables (2.1), we now define

$$v(t) = e^{-\kappa t}z(t) = e^{\kappa(1-t)}\frac{\sinh((-L + \kappa I)t)}{\sinh(-L + \kappa I)}w. \tag{2.5}$$

In the following theorem it is shown that $v(t)$ defined by (2.5) is a good approximation of any solution of (1.2).

THEOREM 2.1 (CF. [2, p.253]). *Let $u(t)$ denote any solution of* (1.2), *and let $v(t)$ be defined by* (2.5). *Then for* $0 \leqslant t \leqslant 1$

$$\|u(t) - v(t)\| \leqslant \delta^t M^{1-t}. \tag{2.6}$$

Proof. Any solution of $u_t = -Lu$ can be written $u(t) = \exp(-Lt)u_0$, for some function u_0 [6, p. 109]. If $u(t)$ is a solution of (1.2), we must have

$$\begin{cases} \|u_0\| \leqslant M, \\ w = \exp(-L)u_0 + \Psi, \\ \|\Psi\| \leqslant \delta. \end{cases} \tag{2.7}$$

We now get

$$\|u(t)-v(t)\| \leqslant \left\| \exp(-Lt) - e^{\kappa(1-t)} \frac{\sinh((-L+\kappa I)t)}{\sinh(-L+\kappa I)} \exp(-L) \right\| \cdot \|u_0\|$$
$$+ e^{\kappa(1-t)} \left\| \frac{\sinh((-L+\kappa I)t)}{\sinh(-L+\kappa I)} \right\| \cdot \|\Psi\|,$$

where the operator norm is defined $\|A\| = \sup\{\|Au\| : \|u\| = 1\}$. We now use (2.7) and the fact that L is selfadjoint and nonnegative to get

$$\|u(t)-v(t)\| \leqslant \sup_{\lambda \geqslant 0} A(\lambda) M + \sup_{\lambda \geqslant 0} B(\lambda)\delta, \tag{2.8}$$

where

$$\begin{cases} A(\lambda) = \left| e^{-\lambda t} - e^{\kappa(1-t)} \dfrac{\sinh((-\lambda+\kappa)t)}{\sinh(-\lambda+\kappa)} e^{-\lambda} \right|, \\ B(\lambda) = e^{\kappa(1-t)} \dfrac{\sinh((-\lambda+\kappa)t)}{\sinh(-\lambda+\kappa)}. \end{cases}$$

It is elementary to show that for $0 \leqslant t \leqslant 1$

$$\sup_{-\infty \leqslant \alpha \leqslant +\infty} \frac{\sinh \alpha t}{\sinh \alpha} = t, \tag{2.9}$$

and therefore

$$\sup_{\lambda \geqslant 0} B(\lambda) = e^{\kappa(1-t)} \cdot t = t(M/\delta)^{1-t}. \tag{2.10}$$

By some simple calculations we find that

$$A(\lambda) = e^{-\kappa t} \frac{\sinh((-\lambda+\kappa)(1-t))}{\sinh(-\lambda+\kappa)},$$

so that by (2.9)

$$\sup_{\lambda \geqslant 0} A(\lambda) = e^{-\kappa t}(1-t) = (\delta/M)^t (1-t). \tag{2.11}$$

Using (2.10) and (2.11) in (2.8), we get the estimate (2.6). Q.E.D.

The backward beam method and the regularization method [5] are related in the sense that the two methods give the same result at $t = 0.5$. This can be seen by the following argument.

In the regularization method an approximate solution of (1.2) is given by $v_r(t) = F(L, t)w$, where the function F is defined by

$$F(\lambda, t) = (\exp(-\lambda) + \mu(t))^{-1} \exp(-\lambda t), \qquad \mu(t) = \frac{\delta}{M} \cdot \frac{1-t}{t}. \tag{2.12}$$

Correspondingly, in the backward beam method we have

$$v(t) = G(L, t)w, \qquad G(\lambda, t) = e^{\kappa(1-t)} \frac{\sinh((-\lambda+\kappa)t)}{\sinh(-\lambda+\kappa)}. \tag{2.13}$$

Therefore, if $F(\lambda, 0.5) = G(\lambda, 0.5)$, for all $\lambda \geq 0$, then $v(0.5) = v_r(0.5)$. By the identity $\sinh 2x = 2 \sinh x \cdot \cosh x$ we get

$$\begin{aligned} G(\lambda, 0.5) &= (M/\delta)^{0.5}(2\cosh((-\lambda + \kappa)/2))^{-1} \\ &= (e^{-\lambda} + \delta/M)^{-1} e^{-\lambda/2} = F(\lambda, 0.5). \end{aligned}$$

It is easy to show that the two methods are different for all other t in the interval $(0, 1)$.

In [2] (2.3) is discretized in time by replacing the derivative z_{tt} by a central difference. Our time discretization is based on the formula (2.5), and the discretization is performed by approximating the exponential function in the hyperbolic sine by a Padé approximation. This is described in the next two sections.

3. Preliminary Error Estimates. We shall now consider an approximation of (2.5), where we replace the exponential function $e^{-\lambda}$ in the hyperbolic sine by a function $f(\lambda)$. In this section we shall discuss this approximation in rather general terms, and it will not be explicitly seen that this corresponds to a time discretization. Here it will be sufficient to distinguish between two classes of approximations of the exponential functions characterized by the following inequalities

$$\text{(3.1)} \qquad \text{(i)} \qquad e^{-\lambda} \leq f(\lambda) \leq 1, \qquad \lambda \geq 0,$$

$$\begin{matrix} \text{(3.2a)} \\ \text{(3.2b)} \end{matrix} \qquad \text{(ii)} \quad \begin{cases} 0 < f(\lambda) \leq e^{-\lambda}, & 0 \leq \lambda \leq \kappa, \\ 0 \leq f(\lambda) \leq 1, & \lambda \geq \kappa. \end{cases}$$

(κ is defined by (2.1).) These two classes correspond to two classes of Padé approximations, as will be seen in the next section.

The approximation of (2.5) is now defined

$$\text{(3.3)} \qquad v_a(t) = e^{\kappa(1-t)} \frac{\sinh((\log f(L) + \kappa I)t)}{\sinh(\log f(L) + \kappa I)} w.$$

Using the spectral representation of L it is easy to see that, under either of the assumptions (3.1) and (3.2), (3.3) is well defined (if $f(\lambda) = 0$ for λ an eigenvalue of L, the expression is to be understood as the limit as λ tends to that eigenvalue).

We now state two theorems which give error estimates for (3.3), when $f(\lambda)$ satisfies (3.1) and (3.2).

THEOREM 3.1. *Let $u(t)$ denote any solution of* (1.2), *let $v_a(t)$ be defined by* (3.3), *and assume that f satisfies* (3.1). *If*

$$\text{(3.4)} \qquad \lambda + \log f(\lambda) \leq (\delta/M)\frac{1}{t}e^{\lambda t} \quad \text{for } 0 \leq \lambda \leq \kappa,$$

then

$$\text{(3.5)} \qquad \|u(t) - v_a(t)\| \leq (t + \max(1, 2(1-t)))\delta^t M^{1-t}.$$

Proof. With the same arguments as in the proof of Theorem 2.1 we get

$$\text{(3.6a)} \qquad \|u(t) - v_a(t)\| \leq \sup_{\lambda \geq 0} A(\lambda) \cdot M + \sup_{\lambda \geq 0} B(\lambda) \cdot \delta,$$

where now

$$\text{(3.6b)}\qquad \begin{cases} A(\lambda) = \left| e^{-\lambda t} - e^{\kappa(1-t)} \dfrac{\sinh((\log f(\lambda) + \kappa)t)}{\sinh(\log f(\lambda) + \kappa)} e^{-\lambda} \right|, \\ B(\lambda) = e^{\kappa(1-t)} \dfrac{\sinh((\log f(\lambda) + \kappa)t)}{\sinh(\log f(\lambda) + \kappa)}. \end{cases}$$

Using (2.9), we immediately get

$$\text{(3.7)}\qquad B(\lambda) \leqslant t(M/\delta)^{1-t}.$$

We then rewrite

$$\begin{aligned} A(\lambda) &= (\delta/M)^t A_1(\lambda) = e^{-\kappa t} A_1(\lambda) \\ &= e^{-\kappa t} \left| e^{(-\lambda+\kappa)t} - \frac{\sinh((\log f(\lambda) + \kappa)t)}{\sinh(\log f(\lambda) + \kappa)} e^{-\lambda+\kappa} \right|. \end{aligned}$$

We now show that $A_1(\lambda) \leqslant 1$ for $\lambda \geqslant \kappa$. Using (2.9), we get

$$0 \leqslant \frac{\sinh((\log f(\lambda) + \kappa)t)}{\sinh(\log f(\lambda) + \kappa)} e^{-\lambda+\kappa} \leqslant t e^{-\lambda+\kappa} \leqslant e^{(-\lambda+\kappa)t},$$

since $0 \leqslant t \leqslant 1$. Therefore, for $\lambda \geqslant \kappa$,

$$\text{(3.8)}\qquad A_1(\lambda) = e^{(-\lambda+\kappa)t} - \frac{\sinh((\log f(\lambda) + \kappa)t)}{\sinh(\log f(\lambda) + \kappa)} e^{-\lambda+\kappa} \leqslant e^{(-\lambda+\kappa)t} \leqslant 1.$$

We then show that, for $0 \leqslant \lambda \leqslant \kappa$, $A_1(\lambda) \leqslant 2(1-t)$. For notational convenience we put

$$\text{(3.9)}\qquad y = \log f(\lambda) + \kappa, \qquad z = -(\lambda + \log f(\lambda)).$$

Note that by the assumption (3.1) $z \leqslant 0$.

Now we can write

$$\begin{aligned} A_1(\lambda) &= \left| \frac{e^{(y+z)t} \sinh y - \sinh(yt) e^{y+z}}{\sinh y} \right| \\ &= \left| \frac{e^{zt}(e^{yt} \sinh y - e^y \sinh(yt)) + (e^{zt} - e^z) e^y \sinh(yt)}{\sinh y} \right| \\ &= e^{zt} \frac{\sinh(y(1-t))}{\sinh y} + (e^{zt} - e^z) e^y \frac{\sinh(yt)}{\sinh y}, \end{aligned}$$

since both terms are positive. By (2.9) and the fact that $z \leqslant 0$, we get

$$A_1(\lambda) \leqslant (1-t) + (e^{zt} - e^z) e^y \cdot t,$$

and then, using the mean value theorem for the second term,

$$A_1(\lambda) \leqslant (1-t) + z(t-1) e^{\theta} e^y t,$$

where θ lies in the interval (z, zt). Thus, since $z \leqslant 0$, we can estimate

$$A_1(\lambda) \leqslant (1-t) + z(t-1) e^{zt} e^y t.$$

We now go back to the original notation (3.9) and use the assumption (3.4)

$$\begin{aligned} A_1(\lambda) &\leqslant (1-t) + (\lambda + \log f(\lambda))(1-t)(f(\lambda))^{-t} e^{-\lambda t} f(\lambda)(M/\delta)t \\ &\leqslant (1-t) + (1-t)(f(\lambda))^{1-t} \leqslant 2(1-t), \end{aligned} \tag{3.10}$$

where the last inequality follows by (3.1).

Combining (3.6), (3.7), (3.8), and (3.10) we get the desired estimate (3.5). Q.E.D.

THEOREM 3.2. *Let $u(t)$ denote an arbitrary solution of* (1.2), *let $v_a(t)$ be defined by* (3.3), *and assume that f satisfies* (3.2). *If*

$$-\lambda - \log f(\lambda) \leqslant (\delta/M)\frac{\log 2e^{\lambda}}{t} \quad \text{for } 0 \leqslant \lambda \leqslant \kappa, \tag{3.11}$$

then

$$\| u(t) - v_a(t) \| \leqslant (t + \max(1, 2(1-t)))\delta^t M^{1-t}.$$

Proof. With the same notation and the arguments of Theorem 3.1 we immediately get

$$B(\lambda) \leqslant t(M/\delta)^{1-t},$$

and, for $\lambda \geqslant \kappa$,

$$A(\lambda) = (\delta/M)^t A_1(\lambda) \leqslant (\delta/M)^t.$$

It remains to be shown that for $0 \leqslant \lambda \leqslant \kappa$

$$A_1(\lambda) \leqslant 2(1-t). \tag{3.12}$$

As in the proof of the preceding theorem, we can write

$$\begin{aligned} A_1(\lambda) &= | a_1(\lambda) - a_2(\lambda) |, \\ a_1(\lambda) &= e^{zt}\frac{\sinh(y(1-t))}{\sinh y}, \\ a_2(\lambda) &= (e^z - e^{zt})e^y \frac{\sinh(yt)}{\sinh y}, \end{aligned}$$

where y and z are defined by (3.9). Note that due to the assumption (3.2) $z \geqslant 0$.

Now (3.12) follows immediately if we can show that $a_i(\lambda) \leqslant 2(1-t)$ for $i = 1, 2$.

Using (2.9) and going back to the original notation (3.9), we get

$$a_1(\lambda) \leqslant \exp((-\lambda - \log f(\lambda))t)(1-t).$$

By (3.11) we have

$$-\lambda - \log f(\lambda) \leqslant (\delta/M)\frac{\log 2e^{\lambda}}{t} \leqslant \frac{\log 2}{t},$$

since $0 \leqslant \lambda \leqslant \kappa = \log(M/\delta)$, and therefore $a_1(\lambda) \leqslant 2(1-t)$.

Using the mean value theorem and (2.9), we can estimate

$$a_2(\lambda) \leqslant e^{z+y} z(1-t)t,$$

and, going back to the original notation (3.9), we now have

$$a_2(\lambda) \leqslant (M/\delta)e^{-\lambda}(-\lambda - \log f(\lambda))(1-t)t.$$

By the assumption (3.11) it follows that

$$a_2(\lambda) \leqslant \log 2(1-t) \leqslant 2(1-t),$$

and the proof is complete. Q.E.D.

It may not be obvious that (3.4) and (3.11) are conditions on how well $f(\lambda)$ approximates $e^{-\lambda}$. However, if $\lambda + \log f(\lambda)$ is small, then $e^{\lambda} f(\lambda)$ is close to 1 and $f(\lambda)$ is a good approximation of $e^{-\lambda}$.

Note that in both theorems $f(\lambda)$ need only be a good approximation of $e^{-\lambda}$ for $0 \leqslant \lambda \leqslant \kappa$. This is true also for the regularization method. The reason is the same in both cases and is indicated in [5, Section 2].

4. A Priori Step Length Estimates for Padé Approximations. The approximations we have had in mind in Section 3 are Padé approximations. In this section we show that the exponential function can be approximated in a way which corresponds to a time discretization. Using the properties of Padé approximations, we then translate the conditions (3.4) and (3.11) of Theorems 3.1 and 3.2, respectively, into a priori conditions on the step length in time for a given Padé approximation.

Since Theorems 3.1 and 3.2 of this paper are completely equivalent to the corresponding theorems in [5], it would be possible to replace this section almost entirely by Section 4 in [5]. For completeness we here give the basic definitions and the theorems without proof.

Assume that the interval [0, 1] has been divided into N equal subintervals, put $k = 1/N$, and assume that $t = nk$ for some integer n. Let $f_{pq}^N(\lambda)$ be defined by

$$f_{pq}^N(\lambda) = \left(\frac{Q_{pq}(k\lambda)}{P_{pq}(k\lambda)} \right)^N, \tag{4.1}$$

where $Q_{pq}(z)/P_{pq}(z)$ is the Padé approximation to e^{-z} defined by

$$Q_{pq}(z) = \sum_{\nu=0}^{q} \frac{(p+q-\nu)!q!}{(p+q)!\nu!(q-\nu)!}(-z)^{\nu}, \tag{4.2a}$$

$$P_{pq}(z) = \sum_{\nu=0}^{p} \frac{(p+q-\nu)!p!}{(p+q)!\nu!(p-\nu)!}z^{\nu}. \tag{4.2b}$$

Two simple approximations of this type are

$$Q_{10}(z)/P_{10}(z) = 1/(1+z),$$
$$Q_{11}(z)/P_{11}(z) = (1-z/2)/(1+z/2),$$

which in connection with ordinary differential equations correspond to the backward Euler and trapezoidal (Crank-Nicolson) methods.

From the following theorems it is seen that the classes of approximations characterized by (3.1) and (3.2) correspond to Padé approximations with q even and q odd, respectively, $q \leqslant p$.

The following quantity will be used in the theorems

$$\sigma_{pq} = \frac{p!q!}{(p+q)!(p+q+1)!}.$$

THEOREM 4.1. *Let $f_{pq}^N(\lambda)$ be defined by* (4.1), (4.2), *and let q be even, $q \leqslant p$.*

(a) *For all $\lambda \geqslant 0$, $e^{-\lambda} \leqslant f_{pq}^N(\lambda) \leqslant 1$.*

(b) *Let $u(t)$ denote any solution of* (1.2), *and put*

$$v_{pq}^N(t) = e^{\kappa(1-t)} \frac{\sinh\left(\left(\log f_{pq}^N(L) + \kappa I\right)t\right)}{\sinh\left(\log f_{pq}^N(L) + \kappa I\right)} w. \tag{4.3}$$

If $N \geqslant \max(1/t, N_1)$, where $N_1 = \kappa(t\sigma_{pq}\kappa e^{\kappa})^{1/(p+q)}$, then

$$\|u(t) - v_{pq}^N(t)\| \leqslant (t + \max(1, 2(1-t)))\delta^t M^{1-t}.$$

THEOREM 4.2. *Let $f_{pq}^N(\lambda)$, $u(t)$ and $v_{pq}^N(t)$ be defined as in Theorem* 4.1. *Assume that q is odd, $q \leqslant p$, N is even and $t = n/N$, where n is an even integer.*

If $N \geqslant \max(1/t, N_2)$, where

$$N_2 = \kappa\left(\frac{2}{\log 2} t\sigma_{pq}\kappa e^{\kappa}\right)^{1/(p+q)},$$

then

$$\text{(a)} \qquad \begin{cases} 0 < f_{pq}^N(\lambda) \leqslant e^{-\lambda} & \textit{for } 0 \leqslant \lambda \leqslant \kappa, \\ 0 \leqslant f_{pq}^N(\lambda) \leqslant 1 & \textit{for } \lambda \geqslant \kappa, \end{cases}$$

and

$$\text{(b)} \qquad \|u(t) - v_{pq}^N(t)\| \leqslant (t + \max(1, 2(1-t)))\delta^t M^{1-t}.$$

In [5] we give a few tables of N_1 and N_2 for different values of p and q. These tables indicate that it is more efficient to use a high order Padé approximation than a low order one like the backward Euler or trapezoidal approximations.

5. Numerical Realization of the Procedure. So far we have paid no attention to the problem of actually computing the backward beam approximation. In fact, at first sight the formula (4.3) may seem to be very unsuitable for numerical computations, since it involves the evaluation of three different functions of the operator L. In this section we shall show how to reduce (4.3), so that the equations we need to solve in the computation of v_{pq}^N involve only quadratic polynomials in L.

For notational convenience we omit the indices pq in this section. (4.3) can be written

$$v^N(t) = G\left(\log f^N(L), t\right)w,$$

where

$$G(\lambda, t) = e^{\kappa(1-t)} \frac{\sinh((\lambda + \kappa)t)}{\sinh(\lambda + \kappa)}.$$

We now set out to rewrite $G(\log f^N(\lambda), t)$. Using the abbreviations $P = P(\lambda/N)$, $Q = Q(\lambda/N)$, and putting $\theta = (\delta/M)^{1/N}$, we get

$$G\left(\log f^N(\lambda), t\right) = \theta^{-(1-t)N} \frac{\left(\theta^{-1}Q/P\right)^{Nt} - \left(\theta P/Q\right)^{Nt}}{\left(\theta^{-1}Q/P\right)^{N} - \left(\theta P/Q\right)^{N}}$$

$$= \left(P/Q\right)^{N(1-t)} \frac{1 - \left(\theta P/Q\right)^{2Nt}}{1 - \left(\theta P/Q\right)^{2N}}.$$

First consider the case when $t = 1/N$. We then get

$$
\begin{aligned}
G(\log f^N(\lambda), t) &= (P/Q)^{N-1}\left[\sum_{j=0}^{N-1}(\theta P/Q)^{2j}\right]^{-1} \\
&= (P/Q)^{N-1}\prod_{j=1}^{N-1}\left[(\theta P/Q)^2 - \theta P/Q 2\cos\frac{\pi j}{N} + 1\right]^{-1} \qquad (5.1)\\
&= (PQ)^{N-1}\prod_{j=1}^{N-1}\left[(\theta P)^2 - \theta PQ 2\cos\frac{\pi j}{N} + Q^2\right]^{-1}.
\end{aligned}
$$

The factors of the denominator are polynomials of degree $2p$ in λ and can be factorized in quadratic factors (since $x^2 - 2\cos(\pi j/N)x + 1$ has no real linear factors, we cannot get any real linear factors in the denominator). Similarly, the numerator in (5.1) can be factorized in linear and quadratic real factors.

Therefore, $v^N(t)$ can be computed by a recursive scheme

$$
(5.2) \qquad \begin{cases} z_0 = w, \\ (\alpha_i L^2 + \beta_i L + \gamma_i I)z_i = S_i(L)z_{i-1}, \qquad i = 1,2,\ldots,p(N-1), \\ v^N(t) = z_{p(N-1)}, \end{cases}
$$

where the S_i are linear or quadratic polynomials.

The arguments can easily be modified to cover the case when $t = n/N$. We get

$$
\begin{aligned}
G(\log f^N(\lambda), t) = (PQ)^{N-n}\prod_{j=1}^{n-1}\left[(\theta P)^2 - \theta PQ 2\cos\frac{\pi j}{n} + Q^2\right] \\
\cdot\prod_{j=1}^{N-1}\left[(\theta P)^2 - \theta PQ 2\cos\frac{\pi j}{N} + Q^2\right]^{-1}.
\end{aligned}
$$

Thus, if L is a second order elliptic operator, $v^N(t)$ can be computed essentially by solving a sequence of $p(N-1)$ fourth order elliptic equations.

In view of (5.2) it is reasonable to take Np as a measure of efficiency of a certain Padé approximation. In [5] we compare the efficiency of different approximation and give evidence that high order approximations are superior to low order ones. This is also confirmed by numerical experiments (see Section 6 of this paper).

Note that in the procedure given in this paper it is not necessary to store simultaneously the solution at all time levels as in the scheme of [2] (it is possible, however, to save storage also in the latter method by using cyclic reduction for solving the large linear system).

Space discretization and the efficient solution of linear algebraic systems corresponding to (5.2) are treated in [4] for the special case when the geometry is rectangular in two dimensions and the coefficients of L are nonconstant but allow separation of variables (see also [3]).

6. Numerical Examples. In the introduction we pointed out that the main importance of the results of this paper would be in connection with problems in two (or more) space dimensions where the geometry is nonrectangular, or the coefficients of L are nonconstant. However, to illustrate the theory we choose instead the simplest example possible, namely the one-dimensional heat equation, because here we know analytic solutions.

Consider the heat equation

$$\begin{cases} u_t = u_{xx}, & 0 \leq x \leq \pi, 0 \leq t \leq 1, \\ u(0, t) = u(\pi, t) = 0. \end{cases}$$

In the Examples 1–3 the solution is taken to be

$$\begin{cases} u(x, t) = c_1 e^{-t} \sin x + c_3 e^{-9t} \sin 3x, \\ c_1 = \sqrt{\frac{2}{\pi}}\, 0.1, \qquad c_3 = \sqrt{\frac{2}{\pi} 0.99}\,. \end{cases} \tag{6.1}$$

Thus we have $M = 1$. The data are chosen as

$$\begin{cases} w(x) = u(x, 1) + \Delta(x), \\ \Delta(x) = c_4 e^{-16} \sin 4x, \qquad c_4 = 8.886\sqrt{\frac{2}{\pi}}\,, \end{cases} \tag{6.2}$$

where the perturbation $\Delta(x)$ has the norm $\delta = \|\Delta(x)\| \approx 10^{-6}$. It is interesting to compare the sizes of the different terms in w:

$$\begin{aligned} 0.1e^{-1} &\approx 3.7 \cdot 10^{-2}, \\ \sqrt{0.99}\, e^{-9} &\approx 1.2 \cdot 10^{-4}, \\ 8.886e^{-16} &\approx 10^{-6}. \end{aligned}$$

In the following we measure the error of an approximate solution $\hat{u}$ as

$$\|u(t) - \hat{u}(t)\|^2 = \frac{\pi}{50} \sum_{i=1}^{49} \left(u(i\pi/50, t) - \hat{u}(i\pi/50, t)\right)^2,$$

which is an approximation of the $L^2[0, \pi]$-norm of the error.

Example 1. We here compare the exact solution (6.1) and the approximation by the backward beam method

$$v(x, t) = \sum_{k=1,3,4} c_k G(k^2, t) e^{-k^2} \sin kx,$$

where G is defined by (2.13). We also give the approximation by the regularization method

$$v_r(x, t) = \sum_{k=1,3,4} c_k F(k^2, t) e^{-k^2} \sin kx,$$

where F is defined by (2.12).

In Figure 6.1 we have plotted the solution and the approximations for four different values of t. The parameter δ/M was chosen equal to 10^{-6} (note that $M = \|u(x, 0)\| = 1$ and $\|\Delta\| = 10^{-6}$). Both methods recognize the perturbation $\Delta(x)$ as noise. This is due to the fact that (for fixed t) the functions $G(\lambda, t)$ and $F(\lambda, t)$ defined by (2.12) and (2.13), respectively, have a maximum at $\lambda = \log(M/\delta)$ (see Figure 6.2 below). Here $\log 10^6 \approx 13.82$.

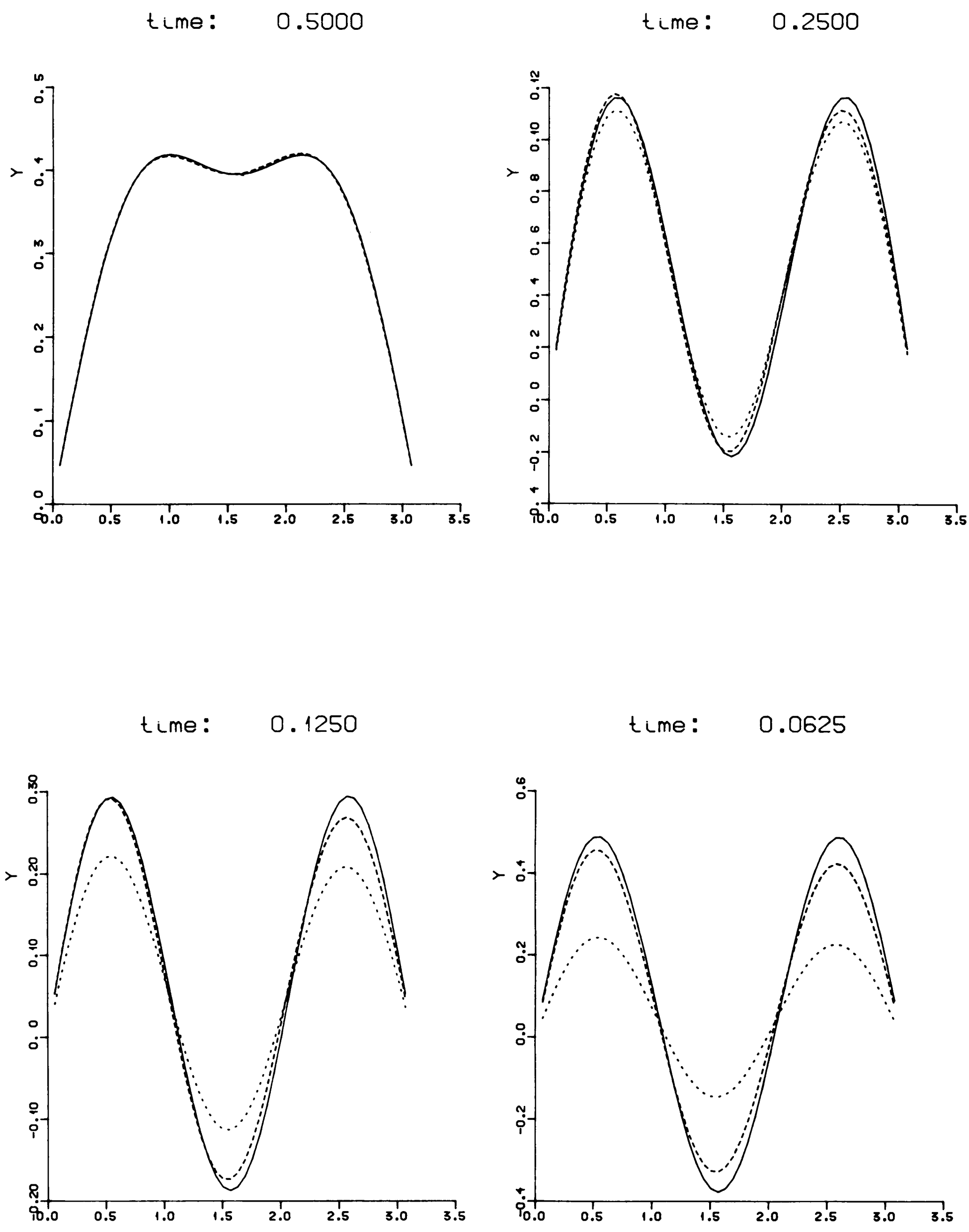

FIGURE 6.1
The backward beam (dotted curve) and the regularization (dashed curve) approximations are plotted together with the correct solution (solid curve). $\delta/M = 10^{-6}$.

Note that for $t = 0.5$ the two methods give identical approximations, but for smaller values of t the regularization method gives better approximations. The reason for this can be seen from Figure 6.2 where we have plotted the functions $G(\lambda, 0.125)$ and $F(\lambda, 0.125)$ defined by (2.13) and (2.12).

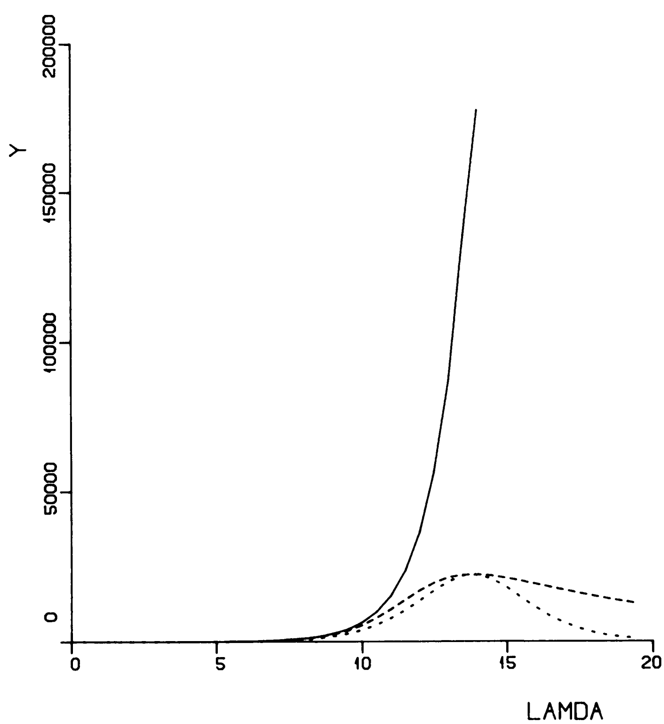

FIGURE 6.2
The functions $G(\lambda, t)$ (*dotted curve*), $F(\lambda, t)$ (*dashed curve*), *and* $\exp(\lambda(1-t))$ (*solid curve*) *are plotted for* $t = 0.125$. $\delta/M = 10^{-6}$.

For $0 \leqslant \lambda \leqslant \log(M/\delta)$ the functions G and F are approximations of $\exp(\lambda(1-t))$. It is easily seen that $G(\lambda, t) \leqslant \exp(\lambda(1-t))$ and $F(\lambda, t) \leqslant \exp(\lambda(1-t))$, but F is a better approximation of the exponential.

In Table 6.3 we give the errors of the approximations for the four values of t. Note that even if the backward beam approximation is not as good as the regularization approximation, the error is still within the theoretical bound.

TABLE 6.3

		Actual error	
t	Theoretical error	Backward beam	Regularization
0.5	10^{-3}	$3.14 \cdot 10^{-4}$	$3.14 \cdot 10^{-4}$
0.25	$3.16 \cdot 10^{-2}$	$1.03 \cdot 10^{-2}$	$6.39 \cdot 10^{-3}$
0.125	0.18	0.098	0.026
0.0625	0.42	0.31	0.066

Example 2. To illustrate the dependence of the value of $\kappa = \log(M/\delta)$ we have computed the backward beam approximation for the problem in Example 1 with $\delta = 10^{-7}$ and 10^{-8}. The results are plotted in Figure 6.4. It is seen that now the noise term in $w(x)$ (6.2) is interpreted as data.

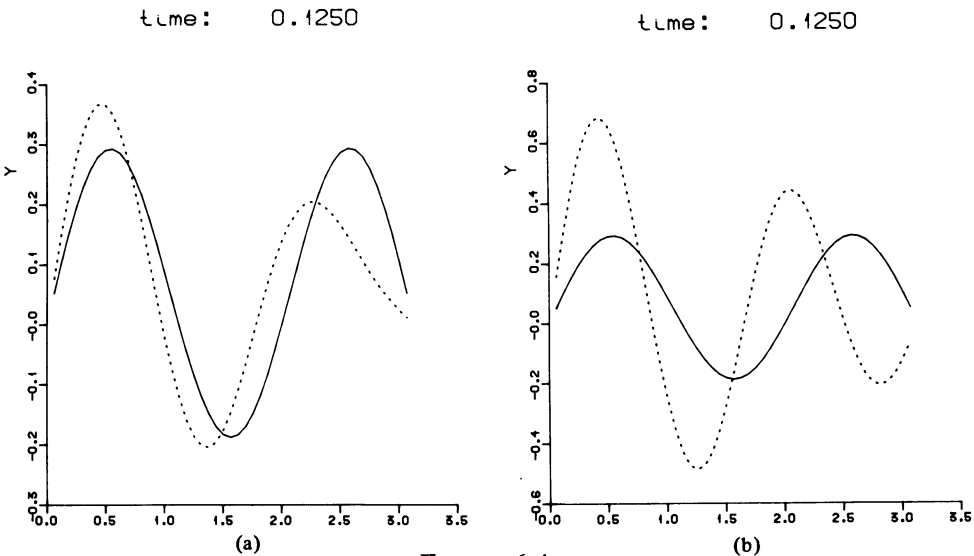

FIGURE 6.4

The backward beam approximation (dotted line) computed at $t = 0.125$ with (a) $\delta/M = 10^{-7}$, *and* (b) $\delta/M = 10^{-8}$. *(Note that the scaling of the vertical axes is different in the two diagrams.) The solid line is the correct solution.*

The explanation for this is that the $G(\lambda, t)$ (see Figure 6.2) has its maximum at $\lambda = \kappa = \log(M/\delta)$; in this case we have $\kappa = \log 10^7 \approx 16.12$ and $\kappa = \log 10^8 \approx 18.42$, which are larger than 16 (see (6.2)).

If we use too small a value of κ, the second term of the solution will be suppressed.

Similar results can be obtained for the regularization method.

Example 3. We consider the same equation and the same data as in Example 1 and study the effect of time discretization by Padé approximation. We have used $\kappa = \log 10^6$ and give the L^2-errors of the time-discrete backward beam and regularization approximations at two time levels. Note that at $t = 0.5$ the two methods are identical.

TABLE 6.5

L^2-error of the backward beam and regularization approximations at $t = 0.5$ for different values of (p, q) and N. The theoretical error estimate (3.5) *is* $1.5 \cdot 10^{-3}$. *The errors in the nondiscrete approximations are both* $3.1 \cdot 10^{-4}$.

(p,q) \ N	2	4	8	10	20	30	40
(1, 1)	$1.1 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$	$8.4 \cdot 10^{-3}$	$4.3 \cdot 10^{-3}$	$8.1 \cdot 10^{-4}$	$3.8 \cdot 10^{-4}$	$3.0 \cdot 10^{-4}$
(2, 2)	$9.7 \cdot 10^{-3}$	$2.1 \cdot 10^{-3}$	$4.1 \cdot 10^{-4}$	$3.5 \cdot 10^{-4}$	$3.2 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$
(3, 3)	$3.0 \cdot 10^{-2}$	$1.6 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$		
(4, 4)	$6.4 \cdot 10^{-4}$	$3.2 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$			
(5, 5)	$4.9 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$					
(6, 6)	$3.8 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$					
(7, 7)	$3.1 \cdot 10^{-4}$						
(8, 8)	$3.1 \cdot 10^{-4}$						
(9, 9)	$3.1 \cdot 10^{-4}$						
(10, 10)	$3.1 \cdot 10^{-4}$						

TABLE 6.6

L^2-error of the regularization approximation at $t = 0.2$ for different values of (p, q) and N. The theoretical error estimate is 0.11. The error in the nondiscrete regularization approximation is $1.1 \cdot 10^{-2}$.

(p, q) \ N	5	10	20	30
(1, 1)	–	$1.1 \cdot 10^{-1}$	$1.7 \cdot 10^{-2}$	$9.9 \cdot 10^{-3}$
(2, 2)	$2.6 \cdot 10^{-2}$	$1.2 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$
(3, 3)	–	$1.1 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$	
(4, 4)	$1.1 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$		
(5, 5)	–			
(6, 6)	$1.1 \cdot 10^{-2}$			

TABLE 6.7

L^2-error of the backward beam approximation at $t = 0.2$ for different values of (p, q) and N. The theoretical error estimate (3.5) is 0.11. The error in the nondiscrete backward beam approximation is $2.5 \cdot 10^{-2}$.

(p, q) \ N	5	10	20	30
(1, 1)	1.2	$6.7 \cdot 10^{-2}$	$7.9 \cdot 10^{-3}$	$1.8 \cdot 10^{-2}$
(2, 2)	$4.1 \cdot 10^{-2}$	$2.6 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$
(3, 3)	$2.4 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	
(4, 4)	$2.5 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$		
(5, 5)	$2.5 \cdot 10^{-2}$			
(6, 6)	$2.5 \cdot 10^{-2}$			

From Tables 6.5–6.7 we see that using a sufficiently high order Padé approximation we can obtain the same accuracy as in the nondiscrete case. If we compare Table 6.5 and Table 4.2 in [5], we also see that for this problem the values of N_1 and N_2 given in Table 4.2 are far too pessimistic. The minimal value of Np as given in Table 5.2 of [5] is 20 ($(p, q) = (10,10)$). The best value in practice is $Np = 8$ ($(p, q) = (4, 4)$).

The numerical results given in Tables 6.5–6.7 support the conclusion that it is safer and more efficient to use a high order approximation and a small value of N than to use a low order approximation and a large N.

Department of Mathematics
University of Linköping
S-581 83 Linköping, Sweden

1. S. AGMON & L. NIRENBERG, "Properties of solutions of ordinary differential equations in Banach space," *Comm. Pure Appl. Math.*, v. 16, 1963, pp. 121–239.

2. B. L. BUZBEE & A. CARASSO, "On the numerical computation of parabolic problems for preceding times," *Math. Comp.*, v. 27, 1973, pp. 237–266.

3. B. L. BUZBEE, "Applications of fast Poisson solvers to A-stable marching procedures for parabolic problems," *SIAM J. Numer. Anal.*, v. 14, 1977, pp. 205–217.

4. L. ELDÉN, "Regularization of the backward solution of parabolic problems," *Inverse and Improperly Posed Problems in Differential Equations* (G. Anger, ed.), Akademie-Verlag, Berlin, 1979.

5. L. ELDÉN, "Time discretization in the backward solution of parabolic equations. I," *Math. Comp.*, v. 39, 1982, pp.

6. A. FRIEDMAN, *Partial Differential Equations*, Holt, Rinehart, and Winston, New York, 1969.

7. V. A. MOROZOV, "On the restoration of functions with guaranteed accuracy," *Numerical Analysis in Fortran*, Moscow Univ. Press, Moscow, 1979, pp. 46–65. (Russian)

8. L. E. PAYNE, *Improperly Posed Problems in Partial Differential Equations*, SIAM, Philadelphia, Pa., 1975.

9. V. N. STRAKHOV, "Solution of incorrectly-posed linear problems in Hilbert space," *Differential Equations*, v. 6, 1970, pp. 1136–1140. (Russian)