UNIVERSITY OF CAMBRIDGE

Numerical Analysis Reports

Numerical solution of the heat equation with nonlocal boundary conditions

Yunkang Liu

DAMTP 1997/NA15
August 1997

Department of Applied Mathematics and Theoretical Physics
Silver Street
Cambridge
England CB3 9EW

Numerical solution of the heat equation with nonlocal boundary conditions

Yunkang Liu

Gonville and Caius College, Cambridge University, Cambridge CB2 1TA, U.K.

Abstract The θ-method for the heat equation with nonlocal boundary conditions is discussed in this paper. The unconditional stability is proved for $\theta \geq 1/2$, subject to a condition that is much weaker than the one assumed in a paper by Ekolin. Due to the nonlocal boundary conditions, the systems of linear equations generated by the θ-method have a coefficient matrix that is tridiagonal except its first and last rows. Three efficient algorithms for solving this kind of linear systems are presented. A numerical example is given to compare their efficiency.

1 Introduction

This paper is concerned with the numerical solution of the heat equation

$$u_t - u_{xx} = f(x,t), \qquad x \in (0,1),\ 0 < t \leq T, \tag{1.1}$$

subject to the nonlocal boundary conditions

$$\begin{cases} u(0,t) = \int_0^1 K_0(x)u(x,t)\,\mathrm{d}x + g_0(t), \\ u(1,t) = \int_0^1 K_1(x)u(x,t)\,\mathrm{d}x + g_1(t), \end{cases} \qquad 0 < t \leq T, \tag{1.2}$$

and the initial condition

$$u(x,0) = u_0(x), \qquad x \in [0,1]. \tag{1.3}$$

The above initial-boundary-value problem arises in the quasi-static theory of thermoelasticity [4, 5]. Similar problems can be found in [1, 2, 11, 15]. For simplicity, we assume that the functions f, K_0, K_1, g_0, g_1 and u_0 are sufficiently smooth. The existence, uniqueness and some analytic properties of the solution of (1.1)–(1.3) have been studied by Day [4, 5] under the assumption that

$$\int_0^1 |K_0(x)|\,\mathrm{d}x < 1, \qquad \int_0^1 |K_1(x)|\,\mathrm{d}x < 1. \tag{1.4}$$

Day's result has been extended to general parabolic equations in several space variables by Friedman [8], Kawohl [10] and others.

The numerical solution of (1.1)–(1.3) and its variations has been considered in several papers. Ekolin [6] proved the convergence of the Crank-Nicolson method subject to the condition

$$\left(\int_0^1 |K_0(x)|^2 \,\mathrm{d}x\right)^{1/2} + \left(\int_0^1 |K_1(x)|^2 \,\mathrm{d}x\right)^{1/2} \leq \frac{\sqrt{3}}{2} \tag{1.5}$$

by using an energy argument. In Ekolin's scheme, the integrals in the nonlocal boundary conditions (1.2) are discretized by the trapezoidal rule. Fairweather and López-Marcos [7] considered the Crank-Nicolson Galerkin method and the extrapolated Crank-Nicolson Galerkin method for the semilinear parabolic equation

$$u_t - (a(x)u_x)_x = F(u, x, t)$$

subject to the nonlocal boundary conditions (1.2) and the initial condition (1.3). Using an energy argument that is different from Ekolin's, they proved the convergences of both methods under the condition that

$$\int_0^1 |K_0(x)|^2 \,\mathrm{d}x < 1, \qquad \int_0^1 |K_1(x)|^2 \,\mathrm{d}x < 1. \tag{1.6}$$

Lin, Xu and Yin [12] studied the backward Euler and a semi-implicit methods for a two dimensional analogue of (1.1)–(1.3). Besides the convergence of the numerical solutions, they also discussed the preservation of some asymptotic features of the exact solution.

One purpose of this paper is to show that (1.5) can be replaced by the condition

$$\int_0^1 |K_0(x)|^2 \,\mathrm{d}x + \int_0^1 |K_1(x)|^2 \,\mathrm{d}x < 2 \tag{1.7}$$

which is much weaker than (1.5). Noting that the θ-method (see, e.g., Morton and Mayers [14]) includes the Crank-Nicolson method ($\theta = 1/2$), we shall prove the convergence of the θ-method for $\theta > 1/2$ under the the assumptions that (1.4) and (1.7) hold. The energy method we used for analysing the θ-method is similar to the one used in Fairweather and López-Marcos [7].

Another purpose of this paper is to formulate efficient algorithms for solving linear systems of equations whose coefficient matrices are tridiagonal except that their first and last rows are full. This kind of linear systems arises when (1.1)–(1.3) is solved by an implicit numerical method such as the θ-method with $\theta > 0$ or the Crank-Nicolson Galerkin method with piecewise linear polynomial interpolation. It also appears in the numerical solution of the heat equation subject to the specification of mass [2, 3]. The numerical experiment in Ekolin [6] shows that if these linear systems of equations are solved by general purpose NAG-library routines, to achieve the same level of accuracy, the backward Euler and the Crank-Nicolson methods for (1.1)–(1.3) consume more CPU time than the forward Euler method does. We shall present in Section 4 of this paper three efficient algorithms. One of them is a straightforward modification of the well-known Thomas algorithm for solving tridiagonal systems of equations. The other two are based on the Sherman–Morrison–Woodbury formula. A numerical example is given at the end of this paper to compare the efficiency of these algorithms.

2 The θ-method

Let $t_n = n\Delta t$ and $x_m = m\Delta x$ for $n = 0, 1, \dots, N$ and $m = 0, 1, \dots, M$, respectively, where $\Delta x = 1/M$ and $\Delta t = 1/N$, $M, N \in \mathbb{Z}^+$, are step sizes in space and time, respectively. The inner product $\langle \cdot, \cdot \rangle$ for $M + 1$ dimensional vectors is defined by

$$\langle u, v \rangle = \Delta x \sum_{k=0}^{M}{}' u_k v_k,$$

where the notation $\sum'$ is such that

$$\sum_{k=0}^{M}{}' w_k = \frac{1}{2} w_0 + \sum_{k=1}^{M-1} w_k + \frac{1}{2} w_M.$$

The ℓ_2 norm $\| \cdot \|$ and the maximum norm $\| \cdot \|_\infty$ are defined, respectively, by

$$\|w\| = \langle w, w \rangle^{1/2},$$

and

$$\|w\|_\infty = \max_{0 \le k \le M} |w_k|.$$

Let u_m^n be our approximation to the exact solution of the initial-boundary-value problem (1.1)–(1.3) at the mesh point (x_m, t_n). Before introducing the θ-method, we define the following difference operators:

$$\begin{aligned}
\mathcal{L} u_m^n &= \Delta_t u_m^n - \Delta_{-x}\Delta_x[(1-\theta)u_m^n + \theta u_m^{n+1}], \\
&\qquad\qquad\qquad m = 1, 2, \dots,\ n = 0, 1, \dots, N-1, \\
\mathcal{L}_0 u^n &= u_0^n - \langle K_0^*, u^n \rangle, \qquad n = 0, 1, \dots, N, \\
\mathcal{L}_1 u^n &= u_M^n - \langle K_1^*, u^n \rangle, \qquad n = 0, 1, \dots, N,
\end{aligned}$$

with $\theta \in [0, 1]$ being a parameter, $u^n = (u_0^n, u_1^n, \dots, u_M^n)^{\mathrm{T}}$ and

$$K_i^* = (K_i(x_0), K_i(x_1), \dots, K_i(x_M))^{\mathrm{T}}, \qquad i = 0, 1.$$

In the definition of $\mathcal{L}$, the following divided difference operators are used:

$$\Delta_t u_m^n = \frac{u_m^{n+1} - u_m^n}{\Delta t}, \qquad \Delta_x u_m^n = \frac{u_{m+1}^n - u_m^n}{\Delta x}, \qquad \Delta_{-x} u_m^n = \frac{u_m^n - u_{m-1}^n}{\Delta x}.$$

The θ-method for (1.1)–(1.3) can be written as

$$\begin{aligned}
&\mathcal{L} u_m^n = f(x_m, (1-\theta)t_n + \theta t_{n+1}), \\
&\qquad\qquad\qquad m = 1, 2, \dots,\ \ n = 0, 1, \dots, N-1, \\
&\begin{cases} \mathcal{L}_0 u^n = g_0(t_n), \\ \mathcal{L}_1 u^n = g_1(t_n), \end{cases} \qquad n = 1, \dots, N, \\
&u_m^0 = u_0(x_m), \qquad m = 0, 1, \dots, M.
\end{aligned}$$

The integrals in the nonlocal boundary conditions (1.2) are discretized by the trapezoidal rule.

The forward Euler method, the Crank-Nicolson method and the backward Euler method correspond, respectively, to the case $\theta = 0, 1/2$ and 1 of the θ-method. Being an explicit method, the forward Euler method is always uniquely solvable. Let us now consider the case $\theta > 0$. Let

$$A = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{M-1} & a_M \\ 1 & \lambda & 1 & & & \\ & 1 & \lambda & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & \lambda & 1 \\ b_0 & b_1 & b_2 & \cdots & b_{M-1} & b_M \end{bmatrix} \tag{2.8}$$

where

$$\lambda = -\frac{1+2\theta\mu}{\theta\mu}$$

with

$$\mu = \frac{\Delta t}{(\Delta x)^2}$$

being the Courant number, and

$$\begin{aligned} &a_0 = 1 - \Delta x K_0(x_0)/2, \quad b_0 = -\Delta x K_1(x_0)/2 \\ &a_m = -\Delta x K_0(x_m), \quad b_m = -\Delta x K_1(x_m), \qquad m = 1, 2, \ldots, M-1, \\ &a_M = -\Delta x K_0(x_M)/2, \quad b_M = 1 - \Delta x K_1(x_M)/2 \end{aligned}$$

Then the numerical solution $u^n = (u_0^n, u_1^n, \ldots, u_M^n)^{\mathrm{T}}$ satisfies the system of equations

$$Au^n = w^n, \tag{2.9}$$

for $n = 1, 2, \ldots, N$, where $w^n = (w_0^n, w_1^n, \ldots, w_N^n)^{\mathrm{T}}$ with

$$\begin{aligned} w_0^n &= g_0(t^n), \\ w_m^n &= \frac{\theta-1}{\theta} u_{m-1}^{n-1} + \frac{2(1-\theta)\mu - 1}{\theta\mu} u_m^{n-1} + \frac{\theta-1}{\theta} u_{m+1}^{n-1} - \frac{\Delta t}{\theta\mu} f(x_m, t_{n-1} + \theta\Delta t), \\ &\qquad\qquad m = 1, 2, \ldots, M-1, \\ w_M^n &= g_1(t^n). \end{aligned}$$

If

$$\Delta x \sum_{m=0}^{M}{}' |K_0(x_m)| < 1, \qquad \Delta x \sum_{m=0}^{M}{}' |K_1(x_m)| < 1,$$

then the matrix A is nonsingular since it is diagonally dominant. Noting that the above inequalities are simply the discretized version of (1.4), we obtain the following trivial result:

Theorem 1 *If* (1.4) *holds then there exists* $h_0 > 0$ *such that the numerical solution of the initial-boundary-value problem (1.1)–(1.3) by the* θ*-method exists and is unique for any* $\Delta x < h_0$.

3 Convergence

Analogous to Theorem 2.4 (for the forward Euler method) and Theorem 2.4 (for the backward Euler method) of Ekolin [6], the following result holds:

Theorem 2 *If* (1.4) *holds and*

$$2(1-\theta)\mu \leq 1 \tag{3.10}$$

then there exists positive constants h_0 and C such that the numerical solution u_m^n of (1.1)–(1.3) satisfies the error estimate

$$|u_m^n - u(x_m, t_n)| \leq C(\Delta x)^2$$

for any $\Delta x < h_0$.

This theorem can be proved without any difficulty by using the maximum principle.

Except the case $\theta = 1$, the preceding result shows the convergence of the θ-method subject to the rather restrictive condition (3.10). However, it is well known that if K_0 and K_1 are zero then the θ-method is convergent for $\theta \geq 1/2$ regardless of the Courant number μ. To show that a similar result holds in the case of nonlocal boundary conditions, we first prove the following stability results by using the energy method.

Lemma 3 *Suppose that both* (1.4) *and* (1.7) *hold. If $\theta \geq 1/2$ then there exist positive constants h_0 and C (both independent of μ) such that*

$$\begin{aligned} \|u^n\|_\infty \quad &\leq \quad C\{\|\Delta_x u^0\| + |u_0^0| + |u_M^0| + |\mathcal{L}_0 u^n| + |\mathcal{L}_1 u^n| \\ &\qquad + \left[\Delta t \sum_{k=0}^{n-1} (\|\mathcal{L} u^k\|^2 + |\mathcal{L}_0 \Delta_t u^k|^2 + |\mathcal{L}_1 \Delta_t u^k|^2)\right]^{1/2} \} \end{aligned} \tag{3.11}$$

for $n = 1, 2, \ldots, N$ and any $\Delta x < h_0$.

Proof Let

$$v_m^n = u_m^n - (1 - x_m)u_0^n - x_m u_M^n, \quad m = 0, 1, \ldots, N, \; n = 0, 1, \ldots, M.$$

Noting that

$$\Delta_{-x}\Delta_x u^n = \Delta_{-x}\Delta_x v^n$$

and

$$v_0^n = v_M^n = 0, \qquad n = 0, 1, \ldots, N,$$

we obtain

$$\begin{aligned} & -\Delta t \langle \Delta_t v^n, \Delta_{-x}\Delta_x[(1-\theta)u^n + \theta u^{n+1}]\rangle \\ = \quad & -\langle v^{n+1} - v^n, \Delta_{-x}\Delta_x[(1-\theta)v^n + \theta v^{n+1}]\rangle \\ = \quad & \langle \Delta_x(v^{n+1} - v^n), \Delta_x[(1-\theta)v^n + \theta v^{n+1}]\rangle \\ = \quad & \frac{1}{2}\left(\|\Delta_x v^{n+1}\|^2 - \|\Delta_x v^n\|^2\right) + \left(\theta - \frac{1}{2}\right)\|\Delta_x \Delta_t v\|^2. \end{aligned}$$

Let $\mathcal{L}u^n = (\mathcal{L}u_0^n, \mathcal{L}u_1^n, ldots, \mathcal{L}u_M^n)^T$. Then

$$\begin{aligned} & \Delta t\langle \Delta_t v^n, \mathcal{L}u^n\rangle \\ \geq \quad & \Delta t\langle \Delta_t v^n, \Delta_t u^n\rangle + \frac{1}{2}\left(\|\Delta_x v^{n+1}\|^2 - \|\Delta_x v^n\|^2\right). \end{aligned} \tag{3.12}$$

Since

$$\langle \Delta_t v^n, \mathcal{L}u^n\rangle \leq \frac{1}{2}\|\Delta_t v^n\|^2 + \frac{1}{2}\|\mathcal{L}u^n\|^2$$

and

$$\langle \Delta_t v^n, \Delta_t u^n\rangle - \frac{1}{2}\|\Delta_t v^n\|^2 = \frac{1}{2}\left[\|\Delta_t u^n\|^2 - \|\Delta_t u^n - \Delta_t v^n\|^2\right]$$

it follows from (3.12) that

$$\|\Delta_x v^{n+1}\|^2 \leq \|\Delta_x v^n\|^2 + \Delta t\left(\|\mathcal{L}u^n\|^2 + \|\Delta_t u^n - \Delta_t v^n\|^2 - \|\Delta_t u^n\|^2\right), \tag{3.13}$$

To estimate $\|\Delta_t u^n - \Delta_t v^n\|$ we write

$$\begin{aligned} \Delta_t u_m^n - \Delta_t v_m^n \quad = \quad & (1-x_m)\,\mathcal{L}_0\Delta_t u^n + x_m\,\mathcal{L}_1\Delta_t u^n \\ & + \Delta x \sum_{i=0}^{M}{}'[(1-x_m)K_0(x_i) + x_m K_1(x_i)]\Delta_t u_i^n. \end{aligned}$$

Since

$$\Delta x \sum_{m=0}^{M}{}'(1-x_m)^2 = \Delta x \sum_{m=0}^{M}{}' x_m^2 = \frac{2+(\Delta x)^2}{6}, \qquad \Delta x \sum_{m=0}^{M}{}' x_m(1-x_m) = \frac{1-(\Delta x)^2}{6},$$

we have

$$\Delta x \sum_{m=0}^{M}{}'[(1-x_m)\,\mathcal{L}_0\Delta_t u^n + x_m\,\mathcal{L}_1\Delta_t u^n]^2 \leq \frac{1}{2}\left(|\mathcal{L}_0\Delta_t u^n|^2 + |\mathcal{L}_1\Delta_t u^n|^2\right)$$

and

$$\begin{aligned} & (\Delta x)^3 \sum_{m=0}^{M}{}'\left\{\sum_{i=0}^{M}{}'[(1-x_m)K_0(x_i) + x_m K_1(x_i)]\Delta_t u_i^n\right\}^2 \\ \leq \quad & (\Delta x)^2\|\Delta_t u^n\|^2 \sum_{i=0}^{M}{}' \sum_{m=0}^{M}{}'[(1-x_m)K_0(x_i) + x_m K_1(x_i)]^2 \\ \leq \quad & \frac{1}{2}(\|K_0\|^2 + \|K_1\|^2)\|\Delta_t u^n\|^2. \end{aligned}$$

Let

$$\delta = \frac{1-\rho_0}{1+\rho_0} > 0,$$

where

$$\begin{aligned} \rho_0 : \quad = \quad & \max\left\{\int_0^1 |K_0(x)|\,\mathrm{d}x, \int_0^1 |K_1(x)|\,\mathrm{d}x,\right. \\ & \left.\frac{1}{2}\left(\int_0^1 |K_0(x)|^2\,\mathrm{d}x + \int_0^1 |K_1(x)|^2\,\mathrm{d}x\right)\right\}. \end{aligned}$$

Using the inequality

$$(a+b)^2 \leq \left(1+\frac{1}{\delta}\right) a^2 + (1+\delta)b^2$$

we obtain

$$\begin{aligned}\|\Delta_t u^n - \Delta_t v^n\|^2 &\leq \left(\frac{1}{2}+\frac{1}{2\delta}\right)(|\,\mathcal{L}_0\Delta_t u^n|^2 + |\,\mathcal{L}_1\Delta_t u^n|^2) \\ &\quad + \frac{1+\delta}{2}(\|K_0\|^2+\|K_1\|^2)\|\Delta_t u^n\|^2.\end{aligned}$$

Let $h_0 > 0$ be such that the inequality

$$\max\left\{\Delta x \sum_{m=0}^{M}{}'|K_0(x_m)|, \Delta x \sum_{m=0}^{M}{}'|K_1(x_m)|, \frac{1}{2}\left(\|K_0\|^2+\|K_1\|^2\right)\right\} < \frac{1+\rho_0}{2}$$

holds for any $\Delta x < h_0$. If follows from (3.13) that

$$\|\Delta_x v^{n+1}\|^2 \leq \|\Delta_x v^n\|^2 + \Delta t\left[\|\,\mathcal{L}u^n\|^2 + \left(\frac{1}{2}+\frac{1}{2\delta}\right)(|\,\mathcal{L}_0\Delta_t u^n|^2 + |\,\mathcal{L}_1\Delta_t u^n|^2)\right]$$

for any $\Delta x < h_0$. Hence,

$$\begin{aligned}\|\Delta_x v^n\|^2 &\leq \|\Delta_x v^0\|^2 + \Delta t\sum_{k=0}^{n-1}\left[\|\,\mathcal{L}u^k\|^2 + \left(\frac{1}{2}+\frac{1}{2\delta}\right)(|\,\mathcal{L}_0\Delta_t u^k|^2 + |\,\mathcal{L}_1\Delta_t u^k|^2)\right], \\ &\qquad n = 1,2,\ldots,N, \end{aligned} \tag{3.14}$$

for $\Delta x < h_0$. Since $v_0^n = 0$, we have

$$\|v^n\|_\infty^2 \leq \Delta_x\|\Delta_x v^n\|^2. \tag{3.15}$$

Noting that

$$\begin{aligned}u_m^n &= v_m^n + (1-x_m)\langle K_0^*, u^n\rangle + x_m\langle K_1^*, u^n\rangle \\ &\quad + (1-x_m)\,\mathcal{L}_0 u^n + x_m\,\mathcal{L}_1 u^n, \qquad m = 1,2,\ldots,\end{aligned}$$

we obtain

$$|u_m^n| \leq |v_m^n| + \frac{1+\rho_0}{2}\|u^n\|_\infty + |\,\mathcal{L}_0 u^n| + |\,\mathcal{L}_1 u^n|, \qquad m = 1,2,\ldots$$

which implies that

$$\|u^n\|_\infty \leq \frac{2}{1-\rho_0}\left(\|v^n\|_\infty + |\,\mathcal{L}_0 u^n| + |\,\mathcal{L}_1 u^n|\right). \tag{3.16}$$

Since

$$\|\Delta_x v^0\| \leq \|\Delta_x u^0\| + \|u_0^0\| + \|u_M^0\|,$$

it follows from (3.14)–(3.16) that there exists a positive constant C such that (3.11) holds. □

Replacing u_m^n in the a priori estimate (3.11) by $u_m^n - u(x_m, t_n)$, we obtain after conducting standard truncation error analysis the following convergence result:

Theorem 4 *Suppose that both* (1.4) *and* (1.7) *hold. If* $\theta \geq 1/2$ *then there exist positive constants* h_0 *and* C *(both independent of* μ*) such that the numerical solution* u_m^n *of (1.1)–(1.3) satisfies the error estimate*

$$|u_m^n - u(x_m, t_n)| \leq \begin{cases} C\left[(\Delta x)^2 + \Delta t\right], & \theta > 1/2, \\ C\left[(\Delta x)^2 + (\Delta t)^2\right], & \theta = 1/2, \end{cases}$$

for any $\Delta x < h_0$, $n\Delta t \leq T$.

Condition (1.7) is much weaker than (1.5). However, numerical experiments indicate that (1.4) is sufficient for the convergence of the numerical solution generated by the θ-method when $\theta \geq 1/2$ without any restriction on the Courant number. The numerical experiments we conducted involve cases where $K_0(x) = K_1(x) = \delta x$ for $\delta \in [\sqrt{3}, 2)$ and $K_0(x) = K_1(x) = \delta x(1-x)$ for $\delta \in [\sqrt{30}, 6)$.

4 Efficient algorithms

When $\theta > 0$ we need to solve repeatedly linear system of equations of the form

$$Au = w,$$

where A is given in (2.8). In explicit form, the preceding system is

$$\begin{cases} \sum_{k=0}^{M} a_k u_k = w_0, \\ u_{m-1} + \lambda u_m + u_{m+1} = w_m, \qquad m = 1, 2, \ldots, M-1, \\ \sum_{k=0}^{M} b_k u_k = w_M. \end{cases} \tag{4.17}$$

In this section, we present three efficient algorithms for solving (4.17).

Our first method is a straightforward modification of the Thomas algorithm. Let

$$d_0 = 0, \qquad e_0 = 1, \qquad f_0 = 0$$

and define the following recurrence relations

$$d_m = -\frac{1}{\lambda + d_{m-1}}, \quad e_m = -\frac{e_{m-1}}{\lambda + d_{m-1}}, \quad f_m = \frac{w_m - f_{m-1}}{\lambda + d_{m-1}}, \qquad m = 1, 2, \ldots, M-1.$$

Then it can be shown by induction from (4.17) that

$$u_m - d_m u_{m+1} = e_m u_0 + f_m, \qquad m = 0, 1, \ldots, M-1. \tag{4.18}$$

Let

$$g_M = 0, \qquad p_M = 1, \qquad q_M = 0$$

and define the following recurrence relations

$$g_m = d_m g_{m+1} + e_m, \quad p_m = d_m p_{m+1}, \quad q_m = d_m q_{m+1} + f_m, \qquad m = M-1, M-2, \ldots, 0.$$

It can be proved by induction from (4.18) that

$$u_m = g_m u_0 + p_m u_M + q_m, \qquad m = 0, 1, \ldots, M. \tag{4.19}$$

Substituting (4.19) into the first and the last equations of (4.17) we find that u_0 and u_M can be obtained by solving the following system

$$u_0 \sum_{k=0}^{M} a_k g_k + u_M \sum_{k=0}^{M} a_k p_k = w_0 - \sum_{k=0}^{M} a_k q_k,$$

$$u_0 \sum_{k=0}^{M} b_k g_k + u_M \sum_{k=0}^{M} b_k p_k = w_M - \sum_{k=0}^{M} b_k q_k.$$

Therefore, the solution of (4.17) can be obtained from (4.19).

Next, we use the Sherman–Morrison–Woodbury formula to design another algorithm for solving (4.17). Let B be a $(M+1)\times(M+1)$ matrix, U and V be $(M+1)\times k$ and $k \times (M+1)$ matrices. Suppose that both B and $I + VB^{-1}U$ are nonsingular. Sometimes referred as the inverse matrix modification formula, the Sherman–Morrison–Woodbury formula reads

$$(B + UV)^{-1} = B^{-1} - B^{-}U(I + VB^{-1}U)^{-1}VB^{-1}. \tag{4.20}$$

An excellent survey of history and applications of (4.20) can be found in Hager [9]. To make use of (4.20) we choose

$$B = \begin{bmatrix} 1 & 0 & & & & \\ 1 & \lambda & 1 & & & \\ & 1 & \lambda & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & \lambda & 1 \\ & & & & 0 & 1 \end{bmatrix},$$

then

$$A = B + UV$$

with

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} a_0 - 1 & a_1 & \cdots & a_{M-1} & a_M \\ b_0 & b_1 & \cdots & b_{M-1} & b_M - 1 \end{bmatrix}.$$

Let y be the solution of

$$By = w$$

and the $(M+1)\times 2$ matrix W be the solution of

$$BW = U.$$

It follows from (4.20) that the solution of (4.17) is

$$u = \left[I - W(I + VW)^{-1}\right] y.$$

Both y and u can be obtained by using the Thomas algorithm since B is a tridiagonal matrix.

A special feature of the matrix A is that, except the first and last rows, its diagonal and its two immediate off-diagonal parts are constant (λ for the diagonal part and 1 for the two immediate off-diagonal parts). To make use of this special feature, we introduce the follow algorithm of Liu and Wu [13] for solving (4.17). Let

$$A_0 = \begin{bmatrix} \alpha & 1 & & & & \\ 1 & \lambda & 1 & & & \\ & 1 & \lambda & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & \lambda & 1 \\ & & & & 1 & \lambda \end{bmatrix},$$

where

$$\alpha = \frac{\lambda - \sqrt{\lambda^2 - 4}}{2}, \quad \text{and} \quad \beta = \frac{\lambda + \sqrt{\lambda^2 - 4}}{2}.$$

are solutions of the quadratic equation

$$x^2 - \lambda x + 1 = 0.$$

Therefore, $A_0 = LU$ with

$$L = \begin{bmatrix} \alpha & & & & & \\ 1 & \alpha & & & & \\ & 1 & \alpha & & & \\ & & \ddots & \ddots & & \\ & & & 1 & \alpha & \\ & & & & 1 & \alpha \end{bmatrix}, \qquad U = \begin{bmatrix} 1 & \beta & & & & \\ & 1 & \beta & & & \\ & & 1 & \beta & & \\ & & & \ddots & \ddots & \\ & & & & 1 & \beta \\ & & & & & 1 \end{bmatrix}.$$

Let y be the solution of

$$A_0 y = w.$$

Because of the LU factorisation of A_0, the above system of equations can be solved via the following recurrence relation

$$\begin{aligned} y_M &= v_M, \\ y_m &= v_m - \beta y_{m+1}, \qquad m = 0, 1, \ldots, M-1, \end{aligned}$$

where v is that solution of

$$Lv = w$$

which is given by the recurrence relation

$$\begin{aligned} v_0 &= \beta w_0, \\ v_m &= \beta(w_m - v_{m-1}), \qquad m = 1, 2, \ldots, M. \end{aligned}$$

The above procedures are stable, since $|\beta| < 1$. Let $z = u - y$, then

$$Az = \begin{bmatrix} w_0 - \sum_{k=0}^{M} a_k y_k \\ 0 \\ \vdots \\ 0 \\ w_M - \sum_{k=0}^{M} b_k y_k \end{bmatrix}$$

Noting that

$$z_{m-1} + \lambda z_m + z_{m+1} = 0, \qquad m = 1, 2, \ldots, M-1$$

we obtain

$$z_m = c_0 \gamma^{M-m} + c_1 \gamma^m, \qquad m = 0, 1, \ldots, M,$$

where

$$\gamma = \frac{-\lambda - \sqrt{\lambda^2 - 4}}{2}.$$

is a solution of the quadratic equation

$$x^2 + \lambda x + 1 = 0,$$

and c_0, c_1 are constants that are determined by the following systems of two equations

$$c_0 \sum_{k=0}^{M} a_k \gamma^{M-k} + c_1 \sum_{k=0}^{M} a_k \gamma^k = w_0 - \sum_{k=0}^{M} a_k y_k,$$

$$c_0 \sum_{k=0}^{M} b_k \gamma^{M-k} + c_1 \sum_{k=0}^{M} b_k \gamma^k = w_M - \sum_{k=0}^{M} b_k y_k.$$

Finally, we obtain the solution of (4.17) by letting $u = y + z$. This algorithm can be derived from the inverse matrix modification formula by choosing B the same as the matrix A_0. However, the procedure formulated above is more intuitive.

To test the above algorithms we use the following example from Day [4]:

$$u_t - u_{xx} = -\mathrm{e}^{-t}\left\{x(x-1) + \frac{\delta}{6(1+\delta)} + 2\right\}, \qquad x \in (0,1),\ 0 < t \le T,$$

$$u(0,t) = -\delta \int_0^1 u(x,t)\,\mathrm{d}x, \qquad t > 0,$$

$$u(1,t) = -\delta \int_0^1 u(x,t)\,\mathrm{d}x, \qquad t > 0,$$

$$u(x,0) = x(x-1) + \frac{\delta}{6(1+\delta)}, \qquad x \in [0,1],$$

where $\delta = 0.0144$. The above example was used by Ekolin [6] to demonstrate that if the structure of the coefficient matrix A is not exploited then the backward Euler and the Crank-Nicolson methods consume more CPU time than the forward Euler method does. Our numerical experiment are performed on a Sun Sparc 10, using MATLAB 4.2. In the four tables below, 'Error' indicates the maximal error

$$\max_{0\leq m\leq M, 1\leq n\leq N} |u_m^n - u(x_m, t_n)|,$$

'Time' represents the CPU time, and $A0, A1, A2$ and $A3$ are, respectively, the standard MATLAB solver and the three algorithm (in order of appearance) described in this section.

Forward Euler method ($\Delta t \approx 0.4 * (\Delta x)^2$)

M	N	Error	Time
32	2561	2.3e-06	50.5
65	10563	5.7e-07	401

Backward Euler method ($\Delta t \approx (\Delta x)^{3/2}$)

M	N	Error	Time (A0)	Time (A1)	Time (A2)	Time (A3)
32	182	5.6e-05	3.98	8.65	8.50	4.22
65	525	1.9e-05	33.3	46.2	49.2	21.7
130	1483	6.8e-06	384	216	285	120

Crank-Nicolson method ($\Delta t \approx \Delta x$)

M	N	Error	Time (A0)	Time (A1)	Time (A2)	Time (A3)
32	32	2.1e-05	0.91	1.31	1.60	0.73
65	66	5.0e-06	4.60	4.92	6.36	2.88
130	131	1.2e-06	37	19	25	11
260	260	3.2e-07	440	76	113	45

Comparing the above four tables, we find that the Crank-Nicolson method is, unsurprisingly, more efficient than the other two. Nevertheless, the forward Euler method still outperforms the backward Euler method. A simple flop count shows that the cost (in terms of numbers of multiplication and division) of solving (4.17) with Gaussian elimination is proportional to M^2, while the cost of solving (4.17) with $A1$, $A2$ and $A3$ are about $14M$, $21M$ and $8M$, respectively. Though $A3$ is the fastest, the other two algorithms can be easily modified to solve more general linear system of equations.

References

[1] J.R. Cannon, *The One-Dimensional Heat Equation*, Encyclopedia of Mathematics and Its Applications, **23**, Addison-Wesley, Reading, Massachusetts, 1984.

[2] J.R. Cannon, S.P. Esteva, Salvador and J. van der Hoek, A Galerkin procedure for the diffusion equation subject to the specification of mass, *SIAM J. Numer. Anal.* **24** (1987), 499–515.

[3] J.R. Cannon, Y. Lin and S. Wang, An implicit finite difference scheme for the diffusion equation subject to mass specification *Internat. J. Engrg. Sci.* **28** (1990), 573–578.

[4] W.A. Day, Extension of a property of the heat equation to linear thermoelasticity and other theories, *Quart. Appl. Math.* **40** (1982), 319–330.

[5] W.A. Day, A decreasing property of solutions of parabolic equations with applications to thermoelasticity, *Quart. Appl. Math.* **41** (1983), 468–475.

[6] G. Ekolin, Finite difference methods for a nonlocal boundary value problem for the heat equation, *BIT* **31** (1991), 245–261.

[7] G. Fairweather and J.C. López-Marcos, Galerkin methods for a semilinear parabolic problem with nonlocal boundary conditions, *Adv. Comput. Math.* **6** (1996), 243–262.

[8] A. Friedman, Monotonic decay of solutions of parabolic equations with nonlocal boundary conditions, *Quart. Appl. Math.* **44** (1986), 401–407.

[9] W.W. Hager, Updating the inverse of a matrix, *SIAM Review* **31** (1989), 221–239.

[10] B. Kawohl, Remark on a paper by W.A. Day on a maximum principle under nonlocal boundary conditions, *Quart. Appl. Math.* **44** (1987), 751–752.

[11] Y. Lin, An inverse problem for a class of quasilinear parabolic equations, *SIAM J. Math. Anal.* **22** (1991), 146–156.

[12] Y. Lin, S. Xu, and H-M. Yin, Finite difference approximations for a class of non-local parabolic equations, *Internat. J. Math. Math. Sci.* **20** (1997), 147–163.

[13] Y. Liu and C.W. Wu, Solution of quasi-tridiagonal system of linear equations, in *Functional analysis, approximation theory and numerical analysis*, pp 259–269, World Sci. Publishing, River Edge, NJ, 1994.

[14] K.W. Morton and D.F. Mayers, *Numerical Solution of Partial Differential Equations*, Cambridge University Press, 1994.

[15] N.I. Yurchuk, Mixed problem with an integral condition for certain parabolic equations, *Diff. Eqs.* **22** (1986), 2117-2126.