Five Ways of Reducing the Crank-Nicolson Oscillations

OLE ØSTERBY
Department of Computer Science, Aarhus University
DK 8000 Aarhus C, Denmark

Abstract - Crank-Nicolson is a popular method for solving parabolic equations because it is unconditionally stable and second order accurate. One drawback of CN is that it responds to jump discontinuities in the initial conditions with oscillations which are weakly damped and therefore may persist for a long time. We present a selection of methods to reduce the amplitude of these oscillations.

1 Introduction

We study finite difference approximations to the simple diffusion equation

$$u_t = bu_{xx} \tag{1}$$

where u_t denotes the partial derivative of u with respect to t and u_{xx} the second partial derivative of u with respect to x, and b is a diffusion coefficient. We use a step size h in the x-direction and k in the t-direction and define $\mu = k/h^2$. The numerical solution at the point $(x,t) = (mh, nk)$ will be denoted v_m^n. Introducing the difference operator

$$\delta^2 v_m^n = \frac{v_{m+1}^n - 2v_m^n + v_{m-1}^n}{h^2} \tag{2}$$

the general θ-method can be written as

$$\frac{v_m^{n+1} - v_m^n}{k} = b(\theta\delta^2 v_m^{n+1} + (1-\theta)\delta^2 v_m^n). \tag{3}$$

Three values of θ are especially important. $\theta = 0$ corresponds to the explicit method, $\theta = 1$ to the implicit method [4] and $\theta = \frac{1}{2}$ to Crank-Nicolson [3].

The two latter are popular since they are unconditionally stable irrespective of the magnitude of $b\mu$, and CN especially so being of second order in time.

One drawback with CN, however, is that the solution has a tendency to oscillate if there are jump discontinuities in the initial condition or between the initial condition and a boundary condition.

Using the implicit method is not the cure. The numerical solution will be smooth but since the method is only first order accurate in time the error will be large unless the time step is reduced considerably, and then we might as well use CN anyway.

2 The growth factor

In order to study the behaviour of the finite difference schemes it is instructive to use the Fourier transform approach as described in [6]. The propagation of the finite difference solution from one time step to the next is governed by the growth factor which for the explicit method is

$$g(\varphi) \;=\; 1 - 4b\mu \sin^2 \frac{\varphi}{2}, \qquad\qquad -\pi \le \varphi \le \pi \tag{4}$$

φ is the parameter in the frequency domain, φ close to 0 corresponds to slowly varying components and φ close to π corresponds to highly oscillatory components of the solution. The latter are present when there are discontinuities in the initial conditions.

From the form of g it is apparent that $g(\varphi)$ is always less than 1. But if $b\mu > \frac{1}{2}$ then $g(\varphi)$ may become less than -1. An absolute value of g greater than 1 implies that the corresponding solution component will be magnified: we have instability. We note that instability appears first (and strongest) for $\varphi \approx \pm\pi$, i.e. for the highly oscillatory components, and that since $g < -1$ these will be propagated as oscillations in time.

For Crank-Nicolson (CN) the growth factor is

$$g(\varphi) \;=\; \frac{1 - 2b\mu \sin^2 \frac{\varphi}{2}}{1 + 2b\mu \sin^2 \frac{\varphi}{2}}, \qquad\qquad -\pi \le \varphi \le \pi \tag{5}$$

It is apparent that $|g(\varphi)| \le 1$ indicating that CN is unconditionally stable irrespective of $b\mu$ and φ. But we note that when $\varphi \approx \pm\pi$ then $g(\varphi) \approx -1$, especially when $b\mu$ is large. This means that the oscillatory components are propagated as weakly damped oscillations in time.

For the implicit method (IM) the growth factor is

$$g(\varphi) \;=\; \frac{1}{1+4b\mu\sin^2\frac{\varphi}{2}}, \qquad\qquad -\pi \le \varphi \le \pi \tag{6}$$

It is easily seen that $0 \le |g(\varphi)| \le 1$ for all $b\mu$ and φ, so the implicit method is unconditionally stable, and furthermore, it will never produce oscillations in time. An interesting point is that $g(\varphi)$ is very small for $\varphi \approx \pm\pi$ and large $b\mu$, so the components for which CN displays the most annoying behaviour are the same components that are damped most by the implicit method.

A device for coping with damped oscillations known from physics is the moving average. If the numerical solution, v_m^n is oscillating in time then we might instead use

$$w_m^n \;=\; \frac{v_m^{n-1}+2v_m^n+v_m^{n-1}}{4}. \tag{7}$$

as an approximation to the solution at $(x,t) = (mh, nk)$. If this is used in connection with CN the growth factor becomes

$$g_{av} \;=\; \frac{g+2+g^{-1}}{4} \;=\; \frac{1}{1-4b^2\mu^2\sin^4\frac{\varphi}{2}} \tag{8}$$

indicating that this method must never be used when $b\mu$ is small but that it has good performance for large values of $b\mu$.

In the following we shall mainly be concerned with the highest frequency component, i.e. $\varphi = \pi$, and shall therefore often refer to the simpler forms of (5), (6) and (8) with $\sin^2\frac{\varphi}{2} = 1$. But also components with $\varphi < \pi$ may give rise to oscillations that need to be damped. When $\varphi < \pi$ then $b\mu\sin^2\frac{\varphi}{2}$ is equal to a $b\mu'$ which is smaller than $b\mu$ so it is equivalent to use the simple form of (5), (6) and (8) and instead consider a range of $b\mu'$-values up to the maximum value $b\mu$.

The following table shows the growth factor for IM and CN and the average (AV) for the highest frequency component for various values of $b\mu$.

Table 1. Growth factors for IM, CN and AV at various $b\mu$.

$b\mu$	IM	CN	AV
0.1	0.7143	0.6667	1.0933
0.5	0.3333	0.0000	–
1	0.2000	−0.3333	−0.3333
10	0.0244	−0.9048	−0.0025
100	0.0025	−0.9900	−0.000025

3 Five ways of reducing the oscillations

In the following we shall assume that the step sizes h and k have been chosen as adequate for the CN solution in the large. Since the oscillations have their origin in the initial condition the cure lies in how to perform the first time step of length k.

1. (AV) The first damping method is the moving average

$$w_m^n = \frac{v_m^{n-1} + 2v_m^n + v_m^{n-1}}{4}. \tag{9}$$

This means that we must carry the CN solution one step further, in order to be able to take the average value at the end point.

2. (IM) As $b\mu$ gets bigger the unwanted oscillations receive less and less damping from CN (cf. Table 1). At $b\mu = 100$ the damping is only 1% per time step. But one initial step with IM will immediately reduce the amplitude of such oscillations by a factor 0.0025. The implicit method is known to be first order accurate in time, but the local error is second order and if we only take one single step with IM before switching to CN then the overall method is still second order accurate in time. Actually it is allowed to take two IM steps (or more) and keep the second order accuracy, but more than two IM steps can not be recommended.

3. (SM) Another suggestion from Table 1 is to choose an initial small time step, k_1, such that the corresponding $b\mu_1$ becomes equal to 0.5 in which case CN itself will eliminate the high frequency component. In practice we should not expect a dramatic effect since there are also other solution components corresponding to values of φ smaller than π and these will not be reduced to zero. Here we might consider taking more than one small step, say M small steps where M could be 5 or 10 or 20. In this way other solution components will be reduced by the appropriate growth factor raised to the M-th power. In order to get back to the 'normal' step size, k, we must take an extra step of length $k - Mk_1$.

4. (Pearson) It is not necessary to aim at a complete annihilation of the oscillations in one step. If the first step is subdivided into M equal steps of length $k_1 = k/M$ as suggested by Pearson [5] then the cumulative damping will be $g^* = g^M$ and a larger value of $b\mu_1$ such as 2 or 5 is acceptable.

5. (EI) Another way of subdividing the first interval is by exponentially increasing subintervals (cf. [1]) where the subintervals are given by $k_i = \beta k_{i-1}, \quad i = 2, \ldots, M$ for some $\beta > 1$ and with $\sum_1^M k_i = k$. This gives a smoother transition from the subintervals to the regular intervals, especially when β is large, i.e. near 2. The **Pearson** method can be viewed as a special case when $\beta = 1$.

The methods have been listed roughly in order of computational complexity:

1. AV requires no additional computing except one extra time step in order to calculate the average value at the end point, and a few extra flops per point to calculate the average.

2. A separate system of equations must be solved for the implicit step.

3. This is also the case when CN is used with a different step size. If M substeps are used within the first ordinary time step this means extra work. And a last substep of length $k - Mk_1$ means that yet another system must be solved. In the constant coefficient case the LU-decomposition can be saved and reused when we proceed with the same step size. Here we shall typically have three different step sizes and therefore three LU-decompositions. Especially when we have more than one space dimension this may mean a considerable amount of extra work.

4. Similar considerations hold for **Pearson** which actually is cheaper than **SM** for the same number, M, of subintervals, since we save the 'uneven' step of length $k - Mk_1$ and the corresponding LU-decomposition.

5. All the M substeps with **EI** are different and we therefore have M LU-decompositions in the first step. When comparing **SM**, **Pearson** and **EI** we must use a smaller value of M with **EI** for the same amount of work.

These considerations are summed up in Table 2.

Table 2. Number of steps and LU-decompositions.

	method	#steps	#LU
1.	**AV**	$N+1$	1
2.	**IM**	N	2
3.	**SM**	$N+M$	3
4.	**P**	$N+M-1$	2
5.	**EI**	$N+M-1$	$M+1$

4 More theory.

In the **Pearson** strategy the first interval of length k is divided into M subintervals of equal length $k_1 = k/M$.

The growth factor for each of these subintervals and for the highest frequency component is therefore according to (5)

$$g = \frac{1 - 2b\mu/M}{1 + 2b\mu/M} \tag{10}$$

and the cumulated damping is given by $g^* = g^M$. If we wish a total damping of say 10^{-p} then

$$|\frac{1-2b\mu/M}{1+2b\mu/M}| \;=\; 10^{-p/M} \tag{11}$$

or

$$|1-\frac{2b\mu}{M}| \;=\; (1+\frac{2b\mu}{M})10^{-p/M}. \tag{12}$$

Only large values of $b\mu$ are interesting for these considerations so we can assume $b\mu > M/2$ and we have

$$\frac{2b\mu}{M}-1 \;=\; (\frac{2b\mu}{M}+1)10^{-p/M} \tag{13}$$

or

$$b\mu \;=\; \frac{M}{2}\cdot\frac{1+10^{-p/M}}{1-10^{-p/M}}. \tag{14}$$

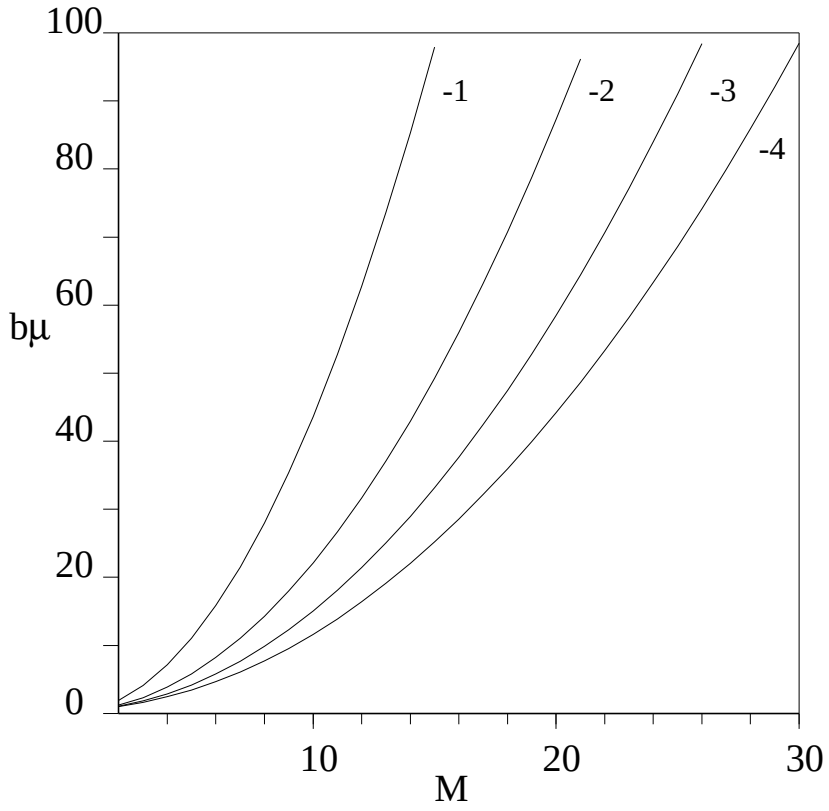

Fig. 1: Damping with **Pearson** as a function of M and $b\mu$.

We thus have a relationship between $b\mu$ and M in order to achieve a given damping of 10^{-p}. This relationship is shown graphically in Fig. 1 for $p = 1, 2, 3, 4$. We note that if $b\mu = 100$ we can achieve a damping of 0.0001 by choosing $M = 30$ or higher.

With exponentially increasing subintervals we choose an initial substep k_1 and a factor $\beta > 1$ such that the following substeps are $k_2 = \beta k_1$, $k_3 = \beta k_2 = \beta^2 k_1$, etc. and such that M substeps equal one ordinary time step:

$$k \;=\; k_1 + \beta k_1 + \beta^2 k_1 + \cdots + \beta^{M-1} k_1 \;=\; k_1 \frac{\beta^M - 1}{\beta - 1} \tag{15}$$

or

$$k_1 \;=\; k \frac{\beta - 1}{\beta^M - 1}. \tag{16}$$

Defining $\mu_i = k_i/h^2$, $i = 1, \ldots, M$ we have

$$\mu_1 \;=\; \mu \frac{\beta - 1}{\beta^M - 1}, \qquad \mu_i \;=\; \beta^{i-1} \mu_1, \quad i = 2, \ldots, M \tag{17}$$

and a total damping of

$$g^* \;=\; \prod_{i=1}^{M} \left| \frac{1 - 2b\mu_i}{1 + 2b\mu_i} \right| \;=\; \prod_{i=1}^{M} \left| \frac{1 - 2\beta^{i-1} b\mu_1}{1 + 2\beta^{i-1} b\mu_1} \right|. \tag{18}$$

For a given value of $b\mu$ we can calculate $g^* = g^*(M, \beta)$. The surface g^* as a function of M and β contains a large number of isolated zeros which occur if one of the involved $b\mu_i$ becomes equal to 0.5. None of these actually show up in Fig. 2 but we at least notice several dips in the surface. The zeros mean that there are several possibilities of achieving complete damping of the highest frequency component. But there are other components (corresponding to $|\varphi| < \pi$) which may give rise to annoying oscillations. These components effectively correspond to (slightly) smaller $b\mu$-values and therefore different $b\mu_i$-values. It is therefore better to focus attention on a 'worst-case' scenario where $b\mu_i$ do not come close to 0.5. The corresponding g^{**} indicate the 'assured' damping in a neighbourhood of β- and $b\mu$-values, such that the observed damping will always be better than or at least as good as g^{**}.

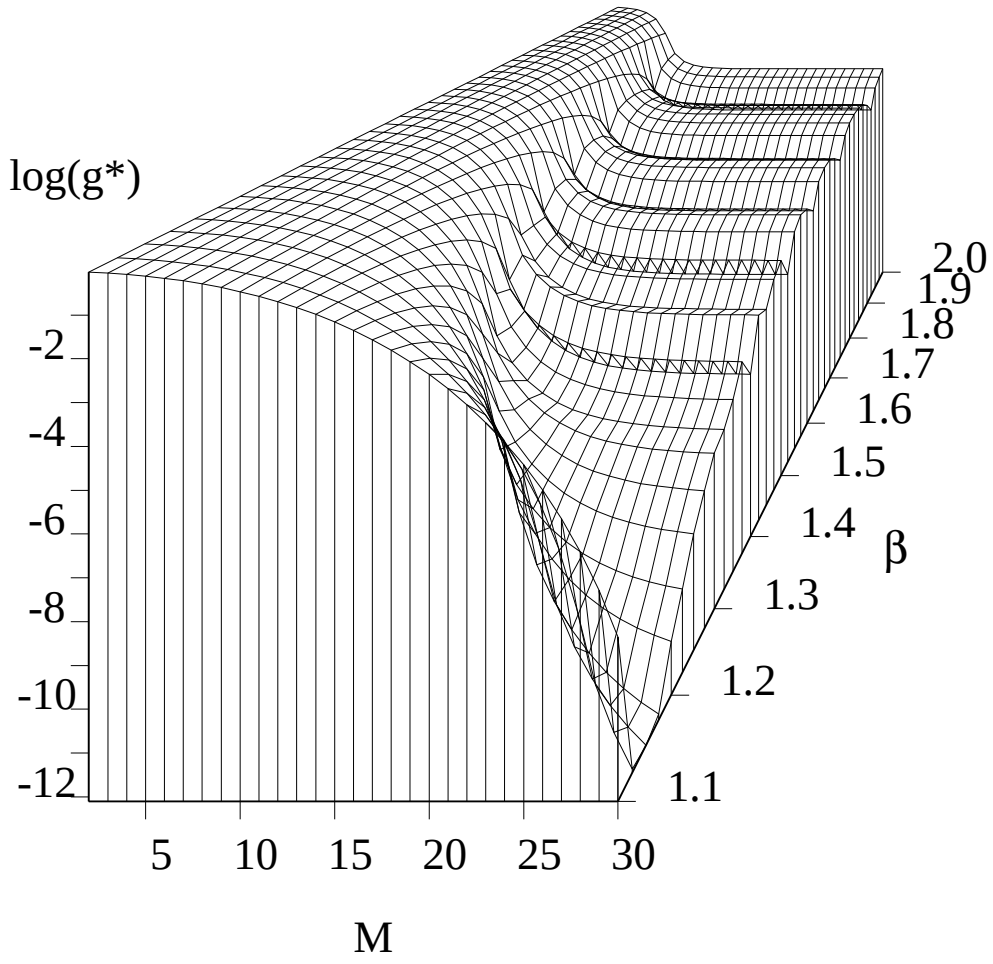

Fig. 2: The damping factor $g^*(M, \beta)$ for **EI** with $b\mu = 100$.

In order to compute g^{**} for a given $b\mu$ and a given M we choose β such that $b\mu_1 < 0.5$, $b\mu_2 = \beta b\mu_1 > 0.5$ and such that

$$\beta b\mu_1 - 0.5 \;=\; 0.5 - b\mu_1 \tag{19}$$

or

$$b\mu_1(\beta + 1) \;=\; 1. \tag{20}$$

Using (17) we get

$$b\mu(\beta^2 - 1) - (\beta^M - 1) \;=\; 0. \tag{21}$$

This equation can be solved numerically for β. With e.g. $b\mu = 100$, $M = 8$, we get $\beta \approx 2.061$ and $g^* = 0.00750$. It is not necessary to require excessive accuracy

in β since the value of g^* exhibits a rather small variation when we are away from the zeros.

Alternatively one could determine β such that

$$\frac{1-2b\mu_1}{1+2b\mu_1} \;=\; g_1 \;=\; -g_2 \;=\; -\frac{1-2\beta b\mu_1}{1+2\beta b\mu_1} \tag{22}$$

or

$$4\beta(b\mu_1)^2 \;=\; 1 \tag{23}$$

or

$$b\mu(\beta-1) - \frac{1}{2}\beta^{-1/2}(\beta^M-1) \;=\; 0. \tag{24}$$

The solution is a slightly different value of β and with a slightly different value of g^*, but the difference is small. In the above example ($b\mu = 100$, $M = 8$) we get $\beta \approx 2.038$, and $g^* = 0.0079$.

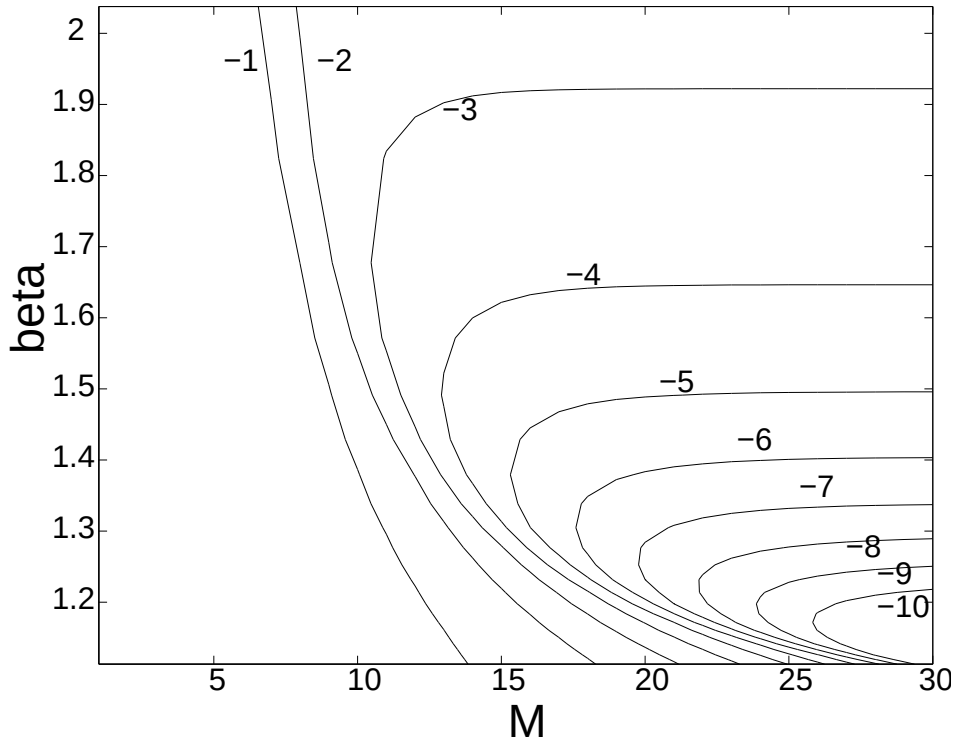

Fig. 3: The contour lines for the assured damping $g^{**}(M,\beta)$ for $b\mu = 100$.

The assured damping is achieved when using this particular value of M. Larger values of M are allowed, but the gain in damping is limited, especially for large values of β.

In the contour plot Fig. 3 we supply values of the assured cumulative damping, g^{**} as a function of M and β for $b\mu = 100$ and for the highest frequency component. Other components correspond to smaller values of $\sin^2 \frac{\varphi}{2}$ or effectively to smaller values of $b\mu$ and this leads to the question how the contour lines of Fig. 3 depend on $b\mu$.

When $b\mu$ and M are not too small then the last '1' in (21) contributes very little and can be left out. If $b\mu$ now is replaced by $b\mu' = b\mu/\beta$ then we get a solution to the modified equation when M is replaced by $M - 1$. In the example we get for $b\mu' = 100/2.061 = 48.52$ that $M = 7$ and $\beta = 2.063$ with $g^* = 0.00752$. So the assured damping is almost the same but it is obtained for a smaller value of M. So the assured damping is mostly dependent on β, assuming that M is chosen large enough.

There is a limit to how much damping can be achieved with a large value of β (around 2 or higher) because successive values of $b\mu_i$ and $b\mu_{i+1} = \beta b\mu_i$ can be very far (on either side) from the optimal value of 0.5, and previous and successive $b\mu_i$ quickly get so far away that they contribute little to the cumulative damping. This is indicated by the almost horisontal contour lines in the upper right of Fig. 3. Smaller values of β allow the $b\mu_i$ to pack closer around 0.5, and successive values are close enough that they can contribute appreciably to g^*. But smaller values of β require larger values of M and thus more work. β should not be chosen too close to 1.0 since this will require a rather large value of M.

We see from Fig. 3 that when $b\mu = 100$ a damping of 0.001 is achieved for $1.5 < \beta < 1.8$ and a value of M of about 12. Smaller values of β are allowed but require larger values of M. A damping of 0.0001 requires $M = 14$ when β is between 1.4 and 1.5. This value can be compared with the value of $M = 30$ required with the **Pearson** strategy for $b\mu = 100$ and the same damping.

When $b\mu$ is larger we would expect the same β-intervals to be valid but M should be chosen larger. The very small damping factors in the lower right part of Fig. 3 will not be noticed in practice because the truncation error will tend to dominate in these cases.

In [2] three of the methods (**IM, Pearson, EI**) have been applied to two problems from electrochemistry where the discontinuity appears between the initial condition and one boundary condition. Here we shall compare the methods on a problem with a jump discontinuity in the initial condition.

5 An example.

If we impose on (1) the initial condition

$$u_0(x) = \begin{cases} 1 & \text{if } |x| < \frac{1}{2} \\ \frac{1}{2} & \text{if } |x| = \frac{1}{2} \\ 0 & \text{if } |x| > \frac{1}{2} \end{cases} \tag{25}$$

then the solution is given by

$$u(x,t) = \frac{1}{2} + 2\sum_{i=0}^{\infty}(-1)^i \frac{\cos \pi(2i+1)x}{\pi(2i+1)} e^{-\pi^2(2i+1)^2 t}. \tag{26}$$

The series in (26) converges reasonably fast for $t > 0$ so only a limited number of terms are necessary. We use the value of (26) at $|x| = 1$ as boundary conditions and solve up to $t = 1$ with $k = 0.01$ and with three different step sizes in the x-direction: $h = 0.02$, 0.01 and 0.005 corresponding to $b\mu = 25$, 100 and 400, respectively. A CN solution with $h = k = 0.01$ at $t = 0.01$ is shown in Fig. 4 on the interval $x \in [-1, 0]$ together with the initial condition. The oscillatory response is strongest at $x = \pm\frac{1}{2} \pm h$ so we shall study the solution at $x = \frac{1}{2} + h$. In Fig. 5 we show the CN-solution at $x = 0.51$ for $t \in [0, 1]$. The oscillations decay with a damping factor of about 0.97 compared to the theoretical factor of 0.99 for the highest frequency component when $b\mu = 100$ (cf. Table 1).

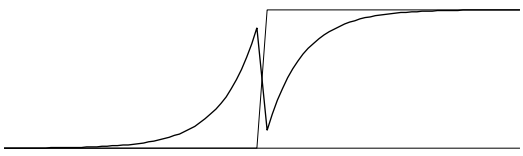

Fig. 4: The CN solution at $t = k = 0.01$ for $x \in [-1, 0]$.
The initial condition is the thin line.

We compare the solutions at $t = 1$ and record the error here. Other intermediate points have been checked, but the end point appears to be representative. In the cases where the oscillations dominate the truncation error, which is signalled by an alternating sign of the error at successive time steps, the recorded error is a good measure of the amplitude of the oscillations. To measure the damping effect of the initial procedure we take the ratio of the measured error to that of the 'pure' CN solution.

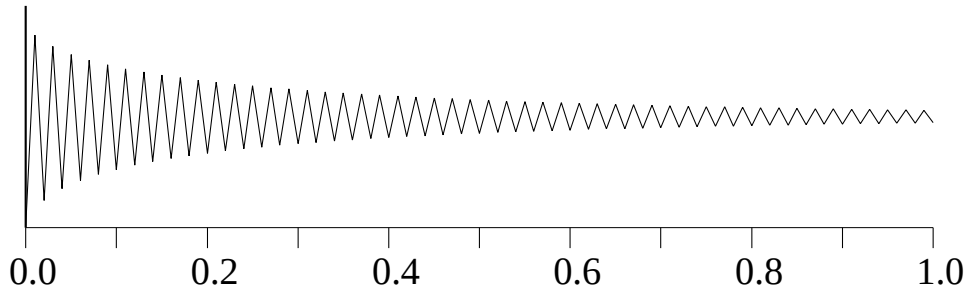

Fig. 5: The CN solution at $x = 0.5 + h = 0.51$ for $t \in [0, 1]$.

In Table 3 we give the damping factors of the various starting procedures at $x = 0.5 + h$ and $t = 1.0$. An asterisk (*) after the number indicates that the oscillations dominate the truncation error such that the error changes sign at successive time steps up to $t = 1.0$. A plus (+) indicate that the oscillations are still visible but with an amplitude smaller than the truncation error, and a minus (−) indicates that the recorded error is monotone decreasing with time. In the latter case the damping factor can only be estimated roughly.

6 Discussion

The truncation error for this example at $(x, t) = (0.5 + h, 1)$ is between 0.000005 and 0.000007 for the three step sizes, except when starting with one or two implicit steps where the error becomes 0.00001 respectively 0.00004.

The amplitude of the (pure) CN oscillations at $(x, t) = (0.5 + h, 1)$ is 0.00039, 0.028 and 0.14 for $h = 0.02$, 0.01 and 0.005, respectively. It is therefore not easy to observe large damping factors when h is large. It is possible (and also more important) to achieve good damping when h is small but generally at a higher value of M when using **SM**, **Pearson** or **EI**.

There is good agreement between Fig. 3 and the results for **EI** (Method 5.) in the column for $b\mu = 100$ in Table 3. One should bear in mind that Fig. 3 reflects the assured damping, and that the observed damping should always be better than or at least as good. When $b\mu = 100$ we actually achieve a damping better than 0.001 with $\beta = 2.0$ contrary to what Fig. 3 predicts. We notice that good damping can be obtained for $\beta = 1.2$ and 1.1 but at the cost of a substantially larger M, in good agreement with Fig. 3.

Comparing to the right and left in Table 3 we observe that the same damping can be obtained at the same value of β but with M increasing (slightly) with $b\mu$.

Good damping can always be achieved with **Pearson** but at a higher value of M. For large values of $b\mu$, M must be considerably larger. Since all the initial substeps are equal this might still be more efficient than **EI** where we also face the intricate problem of selecting an optimal β and M.

Starting with small steps (such that $b\mu = 0.5$) gives results comparable to **Pearson**. M increases with $b\mu$ but is generally smaller than with **Pearson** for the same damping.

IM gives good damping considering the computational cost, especially with two steps. A drawback might be the larger truncation error.

The averaging method actually appears to be most economic giving good damping at a small cost.

The first two damping methods (**AV** and **IM**) are very economic but limited in performance. For the other three methods we can increase M and thus by investing more computer time achieve as good damping as we want.

The first two methods (**AV** and **IM**) perform better with increasing $b\mu$ and as a consequence perform worse for other than the highest frequency component. The actual solution function is a combination of many frequencies and therefore the observed damping is not as good as predicted by the theoretical formulae (6) and (8) with $\varphi = \pi$. On the contrary, **Pearson** and **EI** perform better with decreasing $b\mu$ and therefore give better damping in practice than in theory.

Method/	β	M	$h = 0.02$ $b\mu = 25$		M	$h = 0.01$ $b\mu = 100$		M	$h = 0.005$ $b\mu = 400$	
1.	**AV**		.0010	+		.0001	+		.00002	+
2.	**IM**	1	.015	+	1	.0050	*	1	.0020	*
		2	.001	−	2	.00004	+	2	.00001	+
3.	**SM**	1	.13	*	5	.015	*	20	.0014	*
		5	.0014	+	10	.0010	*	40	.00011	*
		10	.0003	−	20	.00002	+	60	.00001	+
4.	**P**	10	.0081	+	20	.0041	*	40	.0021	*
		12	.0013	+	25	.00047	*	50	.00018	*
		14	.0003	−	30	.00002	+	60	.000006	+
5.	2.0	5	.047	*	9	.00073	*	10	.00043	*
		6	.0060	+	10	.00052	*	12	.000036	+
		7	.0030	+	11	.00025	+	14	.000006	+
	1.7	6	.020	*	9	.00050	*	13	.00013	*
		7	.0016	+	10	.00015	*	14	.000040	+
		8	.0007	+	11	.00009	+	15	.000028	+
	1.4	7	.034	*	10	.027	*	16	.00035	*
		8	.0024	+	12	.00076	*	17	.000010	+
		9	.0003	−	14	.00001	+	18	.000003	+
	1.2	8	.037	*	16	.0012	*	24	.00020	*
		9	.0074	+	17	.00021	+	25	.000030	+
		10	.0016	+	18	.00004	+	26	.000004	+
	1.1	8	.052	*	20	.00090	*	34	.00014	*
		10	.0054	+	22	.00011	+	36	.000015	+
		12	.0006	+	24	.00002	+	38	.000002	+
		13	.0003	−	26	.00001	−	40	.000001	−

Table 3. Damping factors for various starting methods.
The markings (*, +, −) indicate that the oscillations respectively dominate, are of the same order, or are dominated by the truncation error.

References

[1] D. Britz and O. Østerby. Some Numerical Investigations of the Stability of Electrochemical Digital Simulation as Affected by First-order Homogeneous Reactions. J. Electroanal. Chem., 368:143–147, 1994.

[2] D. Britz, O. Østerby and J. Strutwolf. Damping of Crank-Nicolson Error Oscillations. Submitted to Computers Chem., 2002.

[3] J. Crank and P. Nicolson. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. Proc. Camb. Phil. Soc., 43:50–67, 1947. Reprinted in Adv. Comput. Math., 6:207–226, 1996.

[4] P. Laasonen. Über eine Methode zur Lösung der Wärmeleitungsgleichung. Acta Math., 89:309–317, 1949.

[5] C. E. Pearson. Impulsive end condition for diffusion equation. Math. Comp., 19:570–576, 1965.

[6] J. C. Strikwerda. Finite Difference Schemes and Partial Differential Equations. Wadsworth and Brooks/Cole, 1989.