Journal of Engineering Mathematics **31:** 319–335, 1997.

Numerical oscillations on nonuniform grids

PAULA DE OLIVEIRA and FERNANDA PATRÍCIO
Department of Mathematics, University of Coimbra, Coimbra, Portugal

Received 18 November 1994; accepted in revised form 10 April 1996

Key words: Boundary-value problems, Finite differences, nonuniform grids, numerical oscillations.

Abstract. In this paper we study a class of numerical methods used to solve two-point boundary-value problems on nonuniform grids. Particular attention is devoted to numerical oscillations which are quantified for different methods. Numerical experiments are also included.

1. Introduction

The purpose of this paper is to study numerical oscillations for a class of numerical methods used to solve two-point boundary-value problems on nonuniform grids. Our contribution gives a theoretical foundation to numerical results obtained earlier by Veldman and Rinzema [1].

Recently, several adaptive methods have been developed to solve Partial Differential Equations whose solution presents sharp spatial transitions. When standard centered finite-difference formulas are generalized to nonuniform grids, the order of the truncation error is, generally, lower than on uniform grids. However, the use of some of these formulas provide very accurate results. This apparently surprising fact suggests that the global-error should have an order of convergence greater than that of the truncation error. Such a phenomenon, which has been called supraconvergence, has received the attention of many authors. As to supraconvergence of numerical methods for boundary-value problems we can mention for example Manteuffel and White in [2]. With the study of supraconvergence it becomes clear that the truncation error does not provide us with a good indicator of the method's accuracy. In the above-mentioned paper, Veldman and Rinzema study two finite-difference discretizations for a two-point boundary-value problem and conclude that, even if both are supraconvergent, – with the same global-error order – they produce very different numerical simulations. More precisely, the formula which has a first-order truncation error gives more accurate numerical results.

These remarks lead us to the conclusion that the truncation and global-error orders do not give enough information on "the quality" of the numerical simulation. If two formulas have the same global-error order, it seems clear that an indicator to distinguish them could be the size of the error constant. The boundedness properties of this constant are related to stability, but a more detailed analysis of its behaviour can give important information on the expected accuracy.

In the present paper, and following this last idea, we study the numerical oscillations of a class of methods, which includes the methods in [1] and [2], for solving a two-point boundary-value problem on nonuniform grids. Our approach furnishes a prediction of the magnitude of the non-physical oscillations, and also a study of the sensitivity of the method to the index of the node where a step change occurs.

The paper is organized as follows. In Section 2 we construct a general class of methods for solving a two-point boundary-value problem on a nonuniform grid. In Section 3 we study the numerical (non-physical) oscillations of different methods of the class. In Section 4 its asymptotic behaviour (relatively to the second-derivative coefficient) is studied. A certain number of numerical examples which illustrate the accuracy of our predictions are also exhibited. Finally in Section 5 some remarks are presented.

2. A class of methods for solving boundary-value problems

We consider the numerical solution of two-point boundary-value problems of type

$$\begin{cases} -\dfrac{\mathrm{d}T}{\mathrm{d}x} + k\dfrac{\mathrm{d}^2T}{\mathrm{d}x^2} = 0, \quad 0 < x < 1, \quad k > 0, \\ T(0) = 0, \quad T(1) = 1, \end{cases} \tag{2.1}$$

through three-point difference schemes defined on a nonuniform mesh $\{x_i\}_{i=0}^{N}$ with

$$0 = x_0 < x_1 < x_2 \ldots < x_{N-1} < x_N = 1. \tag{2.2}$$

Let

$$h_{j+1} = x_{j+1} - x_j, \quad j = 0, \ldots, N-1, \tag{2.3}$$

$$h = \max_{j=0,\ldots,N-1} h_{j+1}. \tag{2.4}$$

To discretize the first derivative in (2.1) we use a first-order three-point formula defined by

$$\frac{\mathrm{d}}{\mathrm{d}x}T(x_j) = \underline{c}_j T_{j-1} + c_j T_j + \overline{c}_j T_{j+1} + O(h) \tag{2.5}$$

for $j = 1, \ldots, N-1$, where T_j stands for an approximation of $T(x_j)$ and

$$\left\{ \underline{c}_j = -\frac{1}{h_j + h_{j+1}} - \frac{c_j h_{j+1}}{h_j + h_j + 1}, \quad \overline{c}_j = \frac{1}{h_j + h_{j+1}} - \frac{c_j h_j}{h_j + h_{j+1}} \right. . \tag{2.6}$$

In what follows formula (2.5) will be represented by $[\underline{c}_j, c_j, \overline{c}_j]$. From (2.6) we can give (2.5) the form

$$\frac{\mathrm{d}}{\mathrm{d}x}T(x_j) = \frac{T_{j+1} - T_{j-1}}{h_j + h_{j+1}} - c_j \frac{h_j T_{j+1} + h_{j+1} T_{j-1} - (h_j + h_{j+1}) T_j}{h_j + h_{j+1}} + O(h), \tag{2.7}$$

which means that a first-order three-point formula can be viewed as a centered difference formula with a certain amount of numerical viscosity. In fact, the second term on the right-hand side of (2.7) is a discretization of $-\frac{1}{2} c_j h_j h_{j+1} \frac{\mathrm{d}^2}{\mathrm{d}x^2} T$, on the nonuniform grid (2.2). For certain choices of the parameter c_j we find discretization formulas already referred to in the literature. In Table 1 we have listed some of these. For a positive c_j, where $c_j = 1/h_j$, we obtain an upwind difference formula U; if $c_j = 0$ a centered difference formula A is obtained. When $c_j = (h_{j+1} - h_j)/(t h_{j+1} h_j)$, we obtain method B, for $t = 1$, and method C, for $t = 2$; both methods are mentioned in [1] and [2].

Table 1. First-order discretization formulas

Formula designation	c_j	Formula	Main term in truncation error
U	$1/h_j$	$\left[-\frac{1}{h_j}, \frac{1}{h_j}, 0\right]$	$\frac{1}{2} h_j \frac{\mathrm{d}^2}{\mathrm{d}x^2} T(x_j)$
A	0	$\left[-\frac{1}{h_j + h_{j+1}}, 0, \frac{1}{h_j + h_{j+1}}\right]$	$\frac{1}{2}(h_{j+1} - h_j) \frac{\mathrm{d}^2}{\mathrm{d}x^2} T(x_j)$
B	$\frac{h_{j+1} - h_j}{h_{j+1} h_j}$	$\left[-\frac{h_{j+1}}{h_j(h_j + h_{j+1})}, \frac{h_{j+1} - h_j}{h_j h_{j+1}}, \frac{h_j}{h_{j+1}(h_{j+1} + h_j)}\right]$	$\frac{1}{6}(h_{j+1} h_j) \frac{\mathrm{d}^3}{\mathrm{d}x^3} T(x_j)$
C	$\frac{h_{j+1} - h_j}{2h_{j+1} h_j}$	$\left[-\frac{1}{2h_j}, \frac{h_{j+1} - h_j}{2h_j h_{j+1}} \frac{1}{2h_{j+1}}\right]$	$\frac{1}{4}(h_{j+1} - h_j) \frac{\mathrm{d}^2}{\mathrm{d}x^2} T(x_j)$

We shall discretize the second-order derivative in (2.1) using the first-order formula

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}x} T(x_j) \;=\; & \frac{h_j T_{j+1} - (h_j + h_{j+1}) T_j + h_{j+1} T_{j-1}}{h_j h_{j+1} h_{j+1/2}} \\ & - \frac{1}{3}(h_{j+1} - h_j) \frac{\mathrm{d}^3}{\mathrm{d}x^3} T(x_j) + O(h^2), \end{aligned} \tag{2.8}$$

where $h_{j+1/2} = (h_j + h_{j+1})/2$. From (2.7) and (2.8) we obtain a class of methods for solving (2.1) of type

$$\begin{cases} -\frac{T_{j+1} - T_{j-1}}{h_j + h_{j+1}} + \left(\frac{c_j}{2} h_j h_{j+1} + k\right) \frac{h_j T_{j+1} - (h_j + h_{j+1}) T_j + h_{j+1} T_{j-1}}{h_j h_{j+1} h_{j+1/2}} = 0, \quad j = 1, \ldots, N-1, \\ T_0 = 0, \\ T_N = 1. \end{cases} \tag{2.9}$$

If we represent $\frac{T_{j+1} - T_j}{h_{j+1}}$ by $\mathrm{D}T_{j+1}$, for $j = 1, \ldots, N-1$, then (2.9) takes the form

$$\begin{cases} [h_{j+1} - (2k + c_j h_j h_{j+1})]\mathrm{D}T_{j+1} + [h_j + (2k + c_j h_j h_{j+1})]\mathrm{D}T_j = 0, \quad j = 1, \ldots, N-1, \\ T_0 = 0, \\ T_N = 1. \end{cases} \tag{2.10}$$

From (2.10) we can study the oscillatory behaviour of the class of methods. In fact we have

$$\mathrm{D}T_{j+1} = -\frac{h_j + (2k + c_j h_j h_{j+1})}{h_{j+1} - (2k + c_j h_j h_{j+1})} \mathrm{D}T_j, \quad j = 1, \ldots, N-1, \tag{2.11}$$

provided that $h_{j+1} - (2k + c_j h_j h_{j+1}) \neq 0, j = 1, \ldots, N-1$. Let

$$a_{j+1} = h_{j+1} - (2k + c_j h_j h_{j+1}), \quad b_j = h_j + (2k + c_j h_j h_{j+1}), \tag{2.12}$$

for $j = 1, \ldots, N-1$. In order to avoid spurious oscillations at $x = x_j$, the condition

$$b_j / a_{j+1} \leq 0, \quad j = 1, \ldots, N-1, \tag{2.13}$$

Table 2. Necessary and sufficient conditions to avoid oscillation at $x = x_j$

Method Designation	a_{j+1}	b_j	Condition (2.13)
U	$-2k$	$h_j + h_{j+1} + 2k$	always satisfied
A	$h_{j+1} - 2k$	$h_j + 2k$	$h_{j+1} \leq 2k, j = 1, \ldots, N-1$
B	$h_j - 2k$	$h_{j+1} + 2k$	$h_j \leq 2k, j = 1, \ldots, N-1$
C	$\dfrac{h_j + h_{j+1} - 4k}{2}$	$\dfrac{h_j + h_{j+1} + 4k}{2}$	$h_j + h_{j+1} \leq 4k, j = 1, \ldots, N-1$

must be verified. For methods listed in Table 1, the particular form of condition (2.13) is indicated in Table 2. We observe that if h_j is constant, methods A, B and C reduce to the standard centered discretization method, and condition (2.13) in this case takes the well-known form $h \leq 2k$. Moreover, method U is the upwind method which clearly is devoid of spurious oscillations. Since (2.10) is a first-order difference equation, with variable coefficients, it can easily be solved giving

$$T_j = Q_j / Q_N, \quad j = 1, \ldots, N-1, \tag{2.14}$$

with

$$Q_j = \left(1 - \frac{b_1}{a_2}\frac{h_2}{h_1} + \frac{b_1 b_2}{a_2 a_3}\frac{h_3}{h_1} + \cdots + (-1)^{j-1}\frac{b_1 b_2 \cdots b_{j-1}}{a_2 a_3 \cdots a_j}\frac{h_j}{h_1}\right). \tag{2.15}$$

3. Study of numerical oscillations

In this Section we will be concerned with the comparative study of numerical oscillations of methods A, B and C. Methods A and C have a first-order truncation error, but it was proved in [2] that the associated global errors are of second order. Method B has a second-order truncation error and it can easily be established that it has a second-order global-error. If we compare numerical results produced by methods A, B and C, we conclude that, for certain nonuniform grids, method A produces very accurate solutions with practically no spurious oscillations. Methods B and C are less accurate and these lead to significant non-physical oscillations. Consequently, the global error and the truncation-error orders do not give enough information as to the "quality" of the simulation, namely the numerical oscillations. In [1] the authors studied methods A and B following an algebraic approach, but their aim was not to quantify the magnitude of numerical oscillations. Here we follow a different approach which furnishes *a priori* estimations of the magnitude of the oscillations produced by the three methods.

Restrictions on the stepsizes

Problem (2.1) has a boundary layer near $x = 1$. This fact suggests that we should use a mesh of decreasing stepsize, that is

$$h_{j+1} \leq h_j, \quad j = 1, \ldots, N-1.$$

Table 3. Signs of coefficients a_j

Method	Restrictions on h and $\overline{h}$	Sign of a_j according to (2.15)
A	$\overline{h} < 2k$	$a_j > 0, \quad j = 1, \ldots, I$
		$a_j < 0, \quad j = I+1, \ldots, N-1$
B	$\overline{h} < 2k$	$a_j > 0, \quad j = 1, \ldots, I+1$
		$a_j < 0, \quad j = I+2, \ldots, N-1$
C	$\overline{h} < 2k$	$a_j > 0, \quad j = 1, \ldots, I+1$
	$h + \overline{h} > 4k$	$a_j < 0, \quad j = I+2, \ldots, N-1$

Table 4. Numerical viscosity coefficient

Method	Viscosity coefficient
U	$k + \dfrac{h_j}{2}$
A	$k + \dfrac{h_j - h_j + 1}{2}$
B	k
C	$k + \left(\dfrac{h_j - h_{j+1}}{4}\right)$

To simplify the presentation and following [1] we will consider a domain [0,1] decomposed in two subdomains [0, x_I] and [x_I, 1], each one of these being discretized with a uniform mesh of stepsize h and $\overline{h}$, respectively. According to Table 2 we will impose

$$\overline{h} \leq 2k, \tag{2.16}$$

which is a condition that guarantees that no numerical oscillation will appear in $[x_I, 1]$ for method A and in $[x_{I+1}, 1]$ for method B. If we assume that $h + \overline{h} \geq 4$ k, method C will present no oscillation in $[x_{I+1}, 1]$. With this restriction on $\overline{h}$ we can easily analyse (Table 3) the signs of coefficients a_j (coefficients b_j are always positive). These signs will be used in the comparative study of numerical oscillations.

Remark – Following the Modified Equation Approach we know that method (2.9), solves exactly the ordinary differential equation, with an infinite number of terms,

$$-\frac{\mathrm{d}T}{\mathrm{d}x} + \left[k + \frac{c_j}{2} h_j h_{j+1} - \tfrac{1}{2}(h_{j+1} - h_j)\right] \frac{\mathrm{d}^2T}{\mathrm{d}x^2}$$
$$+ \left[\tfrac{1}{3}\left(k + \frac{c_j}{2} h_j h_{j+1}\right)(h_{j+1} - h_j) - \tfrac{1}{6}\frac{h_{j+1}^3 + h_j^3}{h_{j+1} + h_j}\right] \frac{\mathrm{d}^3T}{\mathrm{d}x^3} + \cdots = 0.$$

In this sense this Equivalent Modified Equation has a viscosity coefficient given in Table 4 for the methods under consideration.

Method U has the largest numerical viscosity. In fact, it is well known that upwind solutions contain a large amount of dissipation and no numerical oscillations. Methods A and C are also dissipative, even if they present a smaller amount of numerical viscosity than method U. Method B is not dissipative. If a uniform mesh is used, methods A, and C are not dissipative.

COMPARATIVE STUDY OF NUMERICAL OSCILLATIONS

We recall that we represented by x_I the common "*changing node*" of the two subdomains in which we decomposed [0,1], that is the node where a stepchange occurs. Let d represent an odd number, with $d \leq I$ and v an even number with $v \leq I$. Using the fact that the coefficients b_j are positive for the three methods and the information in Table 3, we easily conclude that

$$Q_d \geq 0 \tag{3.1}$$

Table 5. Behaviour of the numerical solution of (2.10)

Method	Behaviour in $[0, x_I]$		Behaviour in $[x_I, 1]$	
	I odd	I even	I odd	I even
A	oscillating	oscillating	monotone increasing	monotone increasing
B	oscillating	oscillating	monotone	monotone
C	oscillating	oscillating	monotone	monotone

and

$$Q_v \leq 0. \tag{3.2}$$

In fact

$$Q_d = 1 + \left[-\frac{b_1}{a_2} + \frac{b_1 b_2}{a_2 a_3}\right] + \left[(-1)^{d-2}\frac{b_1 b_2 \dots b_{d-2}}{a_2 a_3 \dots a_{d-1}} + (-1)^{d-1}\frac{b_1 b_2 \dots b_{d-1}}{a_2 a_3 \dots a_d}\right], \tag{3.3}$$

where each one of the terms in brackets is positive.

On the other hand, Q_v can be written as

$$Q_v = \left[1 - \frac{b_1}{a_2}\right] + \cdots + \left[\frac{b_1 b_2 \dots b_{v-2}}{a_2 a_3 \dots a_{v-1}}\right] - \left[\frac{b_1 b_2 \dots b_{v-1}}{a_2 a_3 \dots a_v}\right] \tag{3.4}$$

and each one of the terms in brackets is negative. From (3.3) and (3.4) we conclude that for $j \leq I$ the numerical solution is alternately positive and negative and, consequently, oscillatory for the three methods.

Let us now examine the signs of Q_j for $j > I$. We have

$$\begin{aligned} Q_j \;=\; Q_I + \Bigg[&(-1)^I \frac{b_1 b_2 \dots b_I}{a_2 a_3 \dots a_{I+1}} \frac{\overline{h}}{h} + \\ &+(-1)^{I+1} \frac{b_1 b_2 \dots b_{I+1}}{a_2 a_3 \dots a_{I+2}} \frac{\overline{h}}{h} + \cdots + (-1)^{j-1} \frac{b_1 b_2 \dots b_{j-1}}{a_2 a_3 \dots a_j} \frac{\overline{h}}{h}\Bigg]. \end{aligned} \tag{3.5}$$

Let us assume that I is odd. We have from (3.1), $Q_I \geq 0$. From Table 3 we conclude that in (3.5) the sum in brackets is positive for method A. As for this method we have $Q_N > 0$, we conclude from (2.14) that the numerical solution is increasing in $[x_I, 1]$, as is the exact solution of (2.1).

If I is even, we have $Q_N < 0$ for method A and $Q_j < 0$, $j \geq I$. The numerical solution produced by A is also increasing in this case. Concerning methods B and C we conclude from (3.3), (3.4), (3.5), and Table 3 that the solution is monotonic in $[x_I, 1]$. We remark, however, that the sign of Q_N being unknown – for methods B and C – we do not know *a priori* if the solution is increasing or decreasing in $[x_I, 1]$. We summarize these observations in Table 5.

We recall that the uniform stepsize in $[x_I, 1]$, $\overline{h}$, is such that $\overline{h} \leq 2k$. If the uniform stepsize in $[0, x_I]$, h, satisfies $h \leq 2k$, we will have no numerical oscillations. Let us assume that $h > 2k$, which means numerical oscillations will appear in $[0, x_I]$. Let us consider the oscillation w_j of the numerical solution in x_j, for $j \leq I$, defined by

$$w_j = |T_j - T_{j-1}| \tag{3.6}$$

that is

$$w_j = \frac{1}{|Q_N|} \frac{b_1 b_2 \dots b_{j-1}}{|a_2||a_3| \dots |a_j|}. \tag{3.7}$$

Let us suppose that we have established, using (2.13), that the numerical solution is oscillatory in x_j. From (3.6) we can then quantify the oscillation. We observe that $w_j \leq w_{j+1}$ for $j \leq I-1$, where I represents the index of the node where the stepchange occurs.

Remark – The existence of oscillations is detected by (2.13), that is by the sign of $\mathrm{D}T_{j+1}/\mathrm{D}T_j$. From (2.11) we have

$$\frac{\mathrm{D}T_{j+1}}{\mathrm{D}T_j} = -\frac{h_j + (2k + c_j h_j h_{j+1})}{h_{j+1} - (2k + c_j h_j h_{j+1})}. \tag{3.8}$$

Other definitions of numerical oscillations could be proposed. If we had defined the oscillation by this last quotient, we would have had for $j \leq I (h_j = h)$

$$\frac{\mathrm{D}T_{j+1}}{\mathrm{D}T_j} = -\frac{h + 2k}{h - 2k}, \quad j \leq I. \tag{3.9}$$

This expression tells us that $\mathrm{D}T_{j+1}/\mathrm{D}T_j$ is constant for a certain k and a certain stepsize h. As we observe numerically that oscillations increase with j, such a definition would not be an interesting one. As in (3.7) we have $j \leq I$, where I represents the index of the changing node, the term $b_1 b_2 \dots b_{j-1}/|a_2||a_3| \dots |a_j|$ is the same for the three methods. To compare the numerical oscillations w_j it is then sufficient to quantify the different values of Q_N for methods A, B and C. In (3.5) let $j = N$. We obtain

$$\begin{aligned} Q_N \ = \ & Q_I + (-1)^I \frac{b_1 b_2 \dots b_{I-1}}{a_2 a_3 \dots a_I} \frac{\overline{h}}{h} \left[\frac{b_I}{a_{I+1}} + \frac{b_I}{a_{I+1}} \left(-\frac{b_{I+1}}{a_{I+2}} \right) + \right. \\ & \left. + \frac{b_I}{a_{I+1}} \left(-\frac{b_{I+1}}{a_{I+2}} \right) \left(-\frac{b_{I+2}}{a_{I+3}} \right) + \cdots + \left(\frac{b_I}{a_{I+1}} \right) \left(-\frac{b_{I+1}}{a_{I+2}} \right) \dots \left(-\frac{b_{N-1}}{a_N} \right) \right]. \end{aligned} \tag{3.10}$$

Since we assumed that $h_j = h, j = 1, \dots, I$, and $h_j = \overline{h}$ for $j = I+1, \dots, N$, we can simplify (3.10), obtaining

$$Q_N = Q_I + (-1)^I \frac{(b_1)^{I-1}}{(a_2)^{I-1}} \cdot \frac{\overline{h}}{h} \left[\sum_{j=0}^{N-I-1} \frac{b_I}{a_{I+1}} \left(-\frac{b_{I+1}}{a_{I+2}} \right) \right]. \tag{3.11}$$

The sum in brackets – which we will represent by $R(c_I)$ in what follows – is different for the three methods, because it depends on the coefficient c_I (see Table 1). This sum is a geometric sum with $N-I$ terms. We note that $-b_j/a_{j+1}$ is constant for $j = I+1, \dots, N-1$. For methods A, B and C the first term of the sum, b_I/a_{I+1}, and its ratio $-b_{I+1}/a_{I+2}$ are listed in Table 6. We note that the ratio, $-b_{I+1}/a_{I+2}$, is positive and the same for the three methods. The first term, b_I/a_{I+1} is negative for method A and positive for methods B and C. Oscillations produced by methods A and B can now be very easily compared. We recall that for A we have $c_I = 0$ and for B, $c_I = (\overline{h} - h)/(h\overline{h})$. Let us assume that $h^2 + \overline{h}^2 \geq 8k^2$. Under this condition on h and $\overline{h}$ we have

$$-\frac{h + 2k}{\overline{h} - 2k} \geq \frac{\overline{h} + 2k}{h - 2k}. \tag{3.12}$$

Table 6. First term and ratio of the geometric sum $R(c_I)$.

Method	First term $\frac{b_I}{a_{I+1}}$	$-\frac{b_{I+1}}{a_{I+2}}$
A	$\dfrac{h+2k}{\overline{h}-2k}$	$-\dfrac{\overline{h}+2k}{\overline{h}-2k}$
B	$\dfrac{\overline{h}+2k}{h-2k}$	$-\dfrac{\overline{h}+2k}{\overline{h}-2k}$
C	$\dfrac{h+\overline{h}+4k}{h+\overline{h}-4k}$	$-\dfrac{\overline{h}+2k}{\overline{h}-2k}$

Observing now that $Q_I > 0$ for I odd and $Q_I < 0$ for I even, we may conclude

$$\left|Q_I + (-1)^I \frac{b_1 b_2 \dots b_{I-1}}{a_2 a_3 \dots a_I} \frac{\overline{h}}{h} R_I(0)\right| > \left|Q_I + (-1)^I \frac{b_1 b_2 \dots b_{I-1}}{a_2 a_3 \dots a_I} \frac{\overline{h}}{h} R_I\left(\frac{\overline{h}-h}{\overline{h}h}\right)\right|. \quad (3.13)$$

The inequality (3.13) means that the modulus of Q_N, associated with method A, is larger than the modulus of Q_N associated with method B and, consequently, that the oscillations of method A are smaller than the oscillations of method B, once $h^2 + \overline{h}^2 \geq 8k^2$ is satisfied. In Figure 1 we present the numerical solution of (2.1) as computed with Method A and Method B for $k = 10^{-2}, h = 1.9 \times 10^{-1}$ and $\overline{h} = 10^{-2}$. Proceeding as before, we can compare oscillations of methods B and C. We observe that

$$\frac{\overline{h}+2k}{h-2k} \leq \frac{h+\overline{h}+4k}{h+\overline{h}-4k},$$

and, consequently,

$$Q_I + (-1)^I \frac{b_1 b_2 \dots b_{I-1}}{a_2 a_3 \dots a_I} \frac{\overline{h}}{h} R_I\left(\frac{\overline{h}-h}{\overline{h}h}\right) \geq Q_I + (-1)^I \frac{b_1 b_2 \dots b_{I-1}}{a_2 a_3 \dots a_I} \frac{\overline{h}}{h} R_I\left(\frac{\overline{h}-h}{2\overline{h}h}\right). \quad (3.14)$$

We remark that the left-hand side of (3.14) represents the quantity Q_N for method B (for method B we have $c_I = \overline{h} - h/h\overline{h}$) and the right-hand side is the quantity Q_N for method C (for method C, we have $c_I = \overline{h} - h/2h\overline{h}$). In order to compare the relative sizes of the oscillations, defined by (3.7), we must know the signs of both members in (3.14). For example, if they are both positive we conclude that B produces smaller oscillations than C (see Figure 2, which corresponds to $N = 10, I = 5, k = 10^{-3}, h = 1.99 \times 10^{-1}, \overline{h} = 10^{-3}$).

When the two parts of (3.14) are both negative, the oscillations produced by B are larger than those produced by C. (see Figure 3, which corresponds to $N = 10, I = 5, k = 10^{-2}, h = 1.96 \times 10^{-1}, \overline{h} = 10^{-2}$).

4. Asymptotic behaviour of numerical oscillations

As we are interested in the coefficient k, with $k \ll 1$, we study, in this section, the asymptotic behaviour of the oscillation when $k \to 0$. We note that we must have $k \neq 0$.

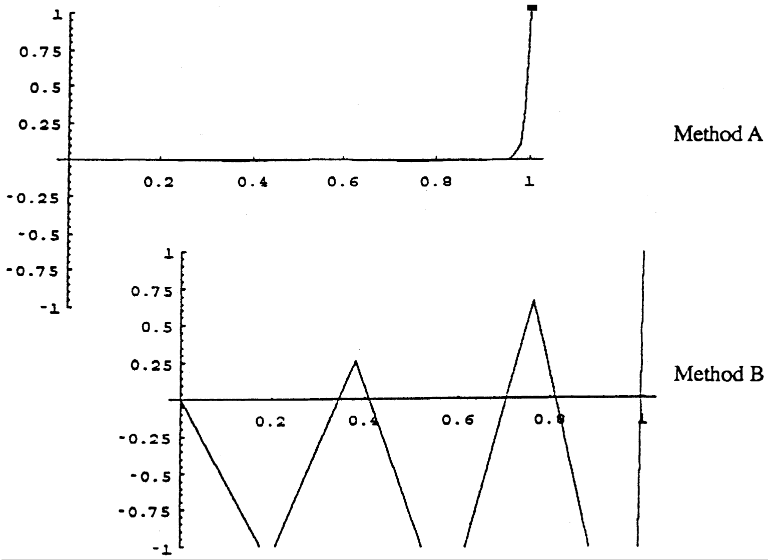

Figure 1. Numerical solution of (2.1) computed with A and B for $k = 10^{-2}$.

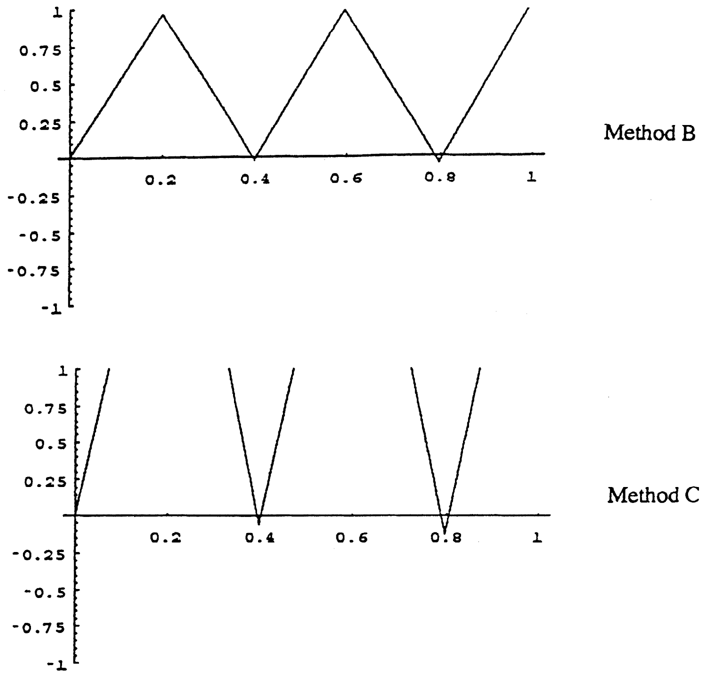

Figure 2. Numerical solutions of (2.1) computed with B and C for $k = 10^{-3}$.

We already assumed that $\overline{h} \leq 2k$. Let us suppose that $\overline{h} = k$. For $j \leq I$, where x_I is the common node of the two subdomains $[0, x_I], [x_I, 1]$, we have

$$w_j = \frac{1}{Q_N} \left[\frac{h + 2k}{h - 2k}\right]^{j-1}. \tag{4.1}$$

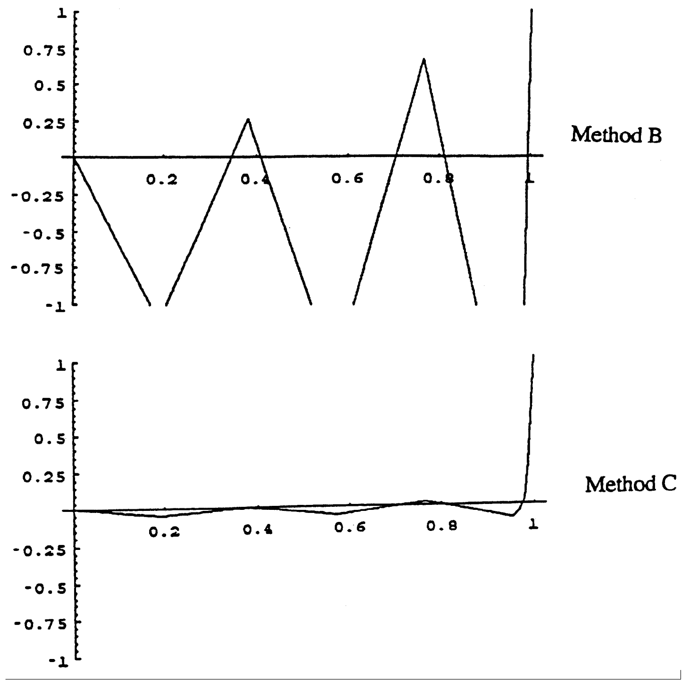

Figure 3. Numerical solutions of (2.1) computed with B and C, for $k = 10^{-2}$.

Method A

From (3.10), Table 3 and Table 6 we have

$$
\begin{aligned}
Q_N \;=\; & Q_I + (-1)^I \left(\frac{h+2k}{h-2k}\right)^{I-1} \frac{\overline{h}}{h} \left(\frac{h+2k}{\overline{h}-2k} + \frac{h+2k}{\overline{h}-2k} \left(-\frac{\overline{h}+2k}{\overline{h}-2k} \right) + \cdots \right. \\
& \left. \cdots + \frac{h+2k}{\overline{h}-2k} \left(-\frac{\overline{h}+2k}{\overline{h}-2k} \right)^{N-I-1} \right)
\end{aligned}
\tag{4.2}
$$

with

$$
Q_I = \left(1 - \left(-\frac{\overline{h}+2k}{\overline{h}-2k} \right)^I \right) \left(\frac{1}{2} - \frac{k}{h} \right). \tag{4.3}
$$

Replacing (4.3) in (4.2) and considering $\overline{h} = k$, we obtain

$$
Q_N = \left[1 - \left(-\frac{h+2k}{h-2k} \right)^I \right] \left(\frac{1}{2} - \frac{k}{h} \right) - (-1)^I \left(\frac{h+2k}{h-2k} \right)^{I-1} \frac{h+2k}{h} \frac{1-3^{N-I}}{2}. \tag{4.4}
$$

Taking limits in (4.1), when $k \to 0$, we easily establish that

$$
\lim_{k \to 0} w_j = \begin{cases} \dfrac{2}{1+3^{N-I}}, & I \text{ odd}, \\[2ex] \dfrac{2}{-1+3^{N-I}}, & I \text{ even}. \end{cases} \tag{4.5}
$$

Table 7. Numerical solution of (2.1) with method A for $k = 10^{-3}$, and $k = 10^{-5}$, respectively.

x_j	Solution with $k = 10^{-3}$	x_j	Solution with $k = 10^{-5}$
0.199	7.491×10^{-3}	0.19999	8.189×10^{-3}
0.398	-1.521×10^{-4}	0.39998	-1.638×10^{-6}
0.597	7.646×10^{-3}	0.59997	8.191×10^{-3}
0.796	-3.104×10^{-4}	0.79996	-3.277×10^{-6}
0.995	7.808×10^{-3}	0.99995	8.193×10^{-3}
0.996	1.601×10^{-2}	0.99996	1.639×10^{-2}
0.997	4.061×10^{-2}	0.99997	4.098×10^{-2}
0.998	1.144×10^{-1}	0.99998	1.148×10^{-1}
0.999	3.358×10^{-1}	0.99999	3.362×10^{-1}

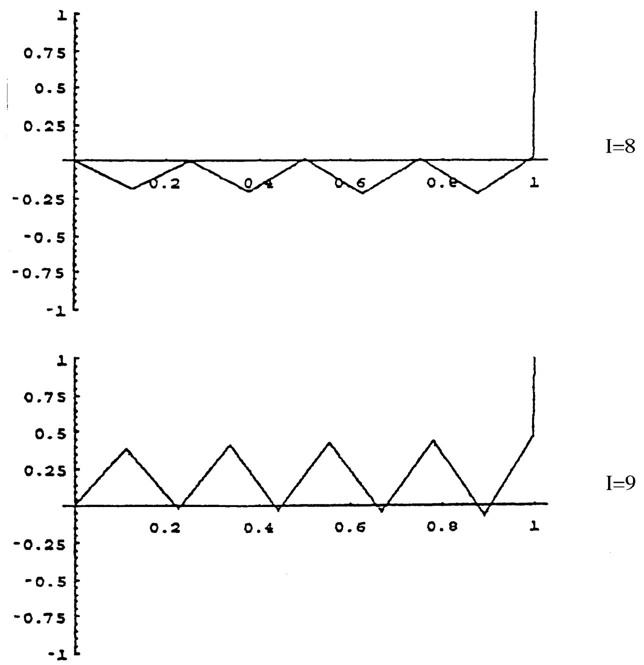

Figure 4. Numerical solutions of (2.1) with method A, for $k = 10^{-3}$ and $N = 10$ and $I = 8$ and $I = 9$.

In Table 7 we present two numerical experiments for $N = 10$, $I = 5$, and respectively $k = 10^{-3}$, $k = 10^{-5}$. From (4.5) we would expect the asymptotic value $w_I \approx 1/122$, which is a good prediction of the numerical oscillations.

In [1] the authors suggested that in convection-dominated problems method A was not very sensitive to the index of the "changing node". This result is confirmed by (4.5). In fact, observing this last expression, we easily see that, with N fixed the more steps of size $\overline{h}$ we consider, that is the smaller is I, then the smaller are the oscillations. In Figure 4 we present two numerical solutions as obtained with method A, for $k = 10^{-3}$, with $N = 10$, and with $I = 8$ and $I = 9$, respectively. From (4.5) we have in the first case $(I = 8), w_I \approx \frac{1}{4}$. In the case $I = 9$ we have $w_I \approx \frac{1}{2}$. Observing that these estimates have been established for $k \to 0$, and that we are using $k = 10^{-3}$, we can conclude that they provide us with good predictions.

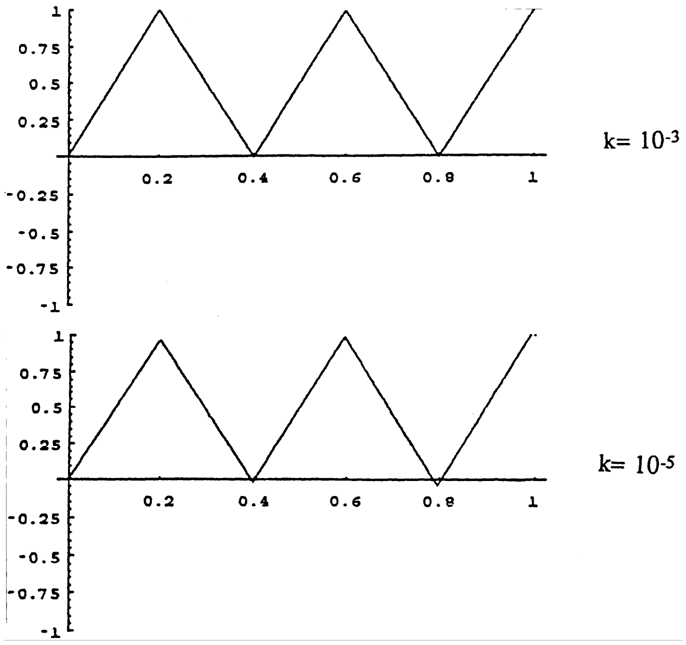

Figure 5. Numerical solutions of (2.1) with method B, for $I = 5$, $N = 10$, $k = 10^{-3}$ and $k = 10^{-5}$, respectively.

Method B

From (3.10), Table 3 and Table 6 we have

$$Q_N = Q_I + (-1)^I \left(\frac{h+2k}{h-2k}\right)^{I-1} \frac{\overline{h}}{h} \left(\frac{\overline{h}+2k}{h-2k} + \frac{\overline{h}+2k}{h-2k}\left(-\frac{\overline{h}+2k}{\overline{h}-2k}\right) + \cdots \right.$$
$$\left. \cdots + \frac{\overline{h}+2k}{h-2k}\left(-\frac{\overline{h}+2k}{\overline{h}-2k}\right)^{N-I-1}\right), \tag{4.6}$$

with Q_I given by (4.3). Considering that, $\overline{h} = k$, in (4.6) and computing the geometric sum in brackets, we have

$$Q_N = \left[1 - \left(-\frac{h+2k}{h-2k}\right)^I\right]\left(\frac{1}{2} - \frac{k}{h}\right) - \frac{1}{2}(-1)^I \left(\frac{h+2k}{h-2k}\right)^{I-1} \frac{3k^2}{h(h-2k)}(1-3^{N-I}). \tag{4.7}$$

Taking limits in (4.1), we obtain

$$\lim_{k \to 0} |\omega_j| = \begin{cases} 1, & I \text{ odd}, \\ \infty, & I \text{ even}. \end{cases} \tag{4.8}$$

In Figure 5 we present two numerical experiments, with method B, for the case I odd, for $k = 10^{-3}$ and $k = 10^{-5}$, respectively. In this experiment $N = 10$ and $I = 5$.

If the parity of the "changing node" x_I is even, it was observed in [1] that the numerical results strongly deteriorate. The result in (4.8) explains this numerical evidence.

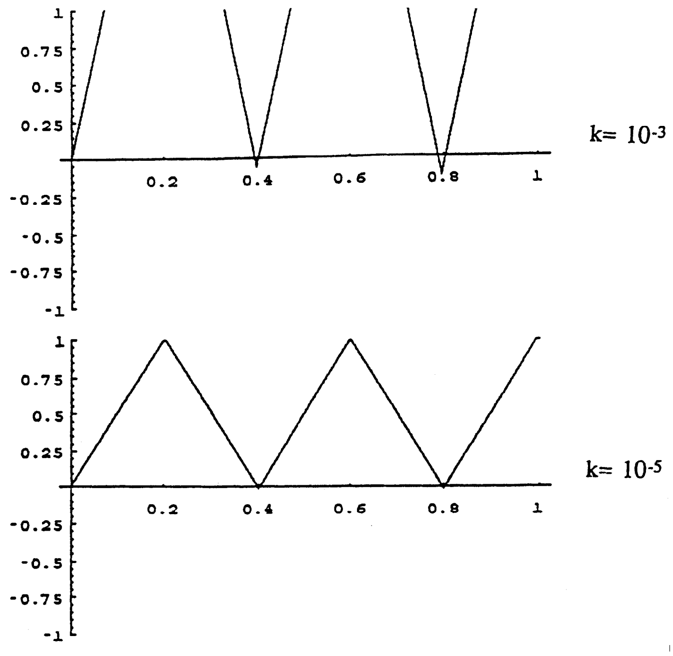

Figure 6. Numerical solutions of (2.1) with method C, for $I = 5$, $N = 10$ and $k = 10^{-3}$, $k = 10^{-5}$, respectively.

Method C

Using (3.10), Table 3 and Table 6, we have

$$Q_N = Q_I + (-1)^I \left(\frac{h+2k}{h-2k}\right)^{I-1} \frac{\overline{h}}{h} \left[\frac{h+\overline{h}+4k}{h+\overline{h}-4k} + \frac{h+\overline{h}+4k}{h+\overline{h}-4k}\left(-\frac{\overline{h}+2k}{\overline{h}-2k}\right) + \cdots \right.$$
$$\left. \cdots + \frac{h+\overline{h}+4k}{h+\overline{h}-4k}\left(-\frac{\overline{h}+2k}{\overline{h}-2k}\right)^{N-I-1}\right], \tag{4.9}$$

with Q_I given by (4.3). Considering that $\overline{h} = k$ in (4.9), and computing the geometric sum in brackets, we obtain

$$Q_N = \left[1 - \left(-\frac{h+2k}{h-2k}\right)^I\right]\left(\frac{1}{2} - \frac{k}{h}\right) - \frac{1}{2}(-1)^I \left(\frac{h+2k}{h-2k}\right)^{I-1} \frac{k}{h}\frac{h+5k}{h-3k}(1-3^{N-I}). \tag{4.10}$$

Taking limits in (4.1), we conclude that

$$\lim_{k\to 0} w_j = \begin{cases} 1, & I \text{ odd} \\ \infty, & I \text{ even.} \end{cases}$$

In Figure 6 we present two numerical experiments, using method C, for $k = 10^{-3}$ and $k = 10^{-5}$, with $I = 5$ and $N = 10$. We note that for $k = 10^{-5}$ methods B and C produce practically the same numerical solution as could be expected from (4.8) and (4.10). For I even the numerical solution exhibits an unstable behaviour.

Table 8. Size of the numerical oscillations when $k \to 0$

Method	I odd	I even
A	$\frac{2}{1+3^{N-I}}$	$\frac{2}{3^{N-I}-1}$
B	1	unbounded
C	1	unbounded

Remark – Let

$$E_{j+1} = T(x_{j+1}) - T_{j+1}, \quad j = 0, 1, \dots, N-1,$$

where $T(x_j)$ represents the solution of (2.1) and T_j the solution of (2.9).

This error satisfies the difference equation

$$\begin{cases} a_{j+1}DE_{j+1} + b_j DE_j = t_j, \quad j = 1, \dots, N-1, \\ E_0 = E_N = 0, \end{cases}$$

where t_j represents the truncation error at $x = x_j$. Proceeding as in section 2, we have

$$E_{j+1} = -\frac{Q_{j+1}}{Q_N} \sum_{i=1}^{N-1} h_{i+1} S_i + \sum_{i=1}^{j} h_{i+1} S_i$$

with

$$S_i = \sum_{l=1}^{i} (-1)^{i-l} \frac{b_{l+1} \dots b_j}{a_{l+2} \dots a_{j+1}} \frac{h_l + h_{l+1}}{h_{l+1} - 2k} t_l.$$

Consequently

$$|E_{j+1}| \leq \left(1 + \frac{|Q_{j+1}|}{|Q_N|}\right) \left(\sum_{i=1}^{N-1} h_{i+1} \sum_{l=1}^{i} \frac{h_l + h_{l+1}}{|h_{l+1} - 2k|} |t_l|\right).$$

Using the truncation errors in Table 1, we conclude that methods A, B and C have a global-error of order two if the constant

$$\left(1 + \frac{|Q_{j+1}|}{|Q_N|}\right), \quad j = 0, 1, \dots, N-1,$$

is bounded. Consequently, methods B and C are clearly unstable when the changing node x_I has an even index. In Table 8 we summarize our conclusions concerning the size of the numerical oscillations for the three methods when $k \to 0$.

To conclude this section, we briefly refer to what happens to the numerical solution when a uniform grid is used. In this case methods A, B and C coincide and from (3.10) we conclude that

$$\lim_{k \to 0} Q_N = \begin{cases} 1, & \text{if } N \text{ is odd} \\ 0, & \text{if } N \text{ is even.} \end{cases}$$

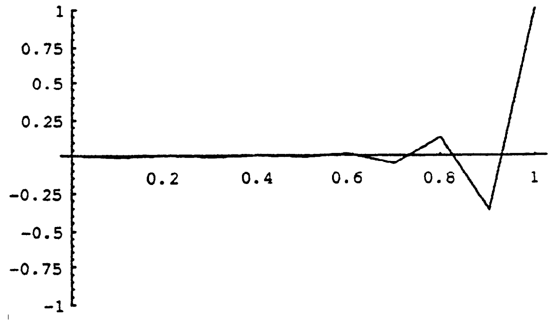

Figure 7. Numerical solution of (2.1) with 4PU ($q = \frac{1}{2}$) on a uniform grid, with $h = 10^{-1}$, $k = 10^{-3}$.

and consequently

$$\lim_{k\to 0} w_j = \begin{cases} 1, & N \text{ is odd} \\ \infty, & N \text{ is even.} \end{cases}$$

As is well known, when N is even, the method is unstable (w_I is unbounded). For N odd, the numerical solution obtained on a uniform grid is analogous to the solutions obtained with methods B and C on a nonuniform grid.

5. Final Remarks

In boundary-value problems such as (2.1), with small viscosity k, the solution exhibits a very sharp profile. If symmetric three-point formulas are used, on a uniform grid, to represent $\mathrm{d}u/\mathrm{d}x$, non-physical oscillations occur; if asymmetric algebraic formulas of low order are used to represent $\mathrm{d}u/\mathrm{d}x$, the smoothness of the numerical solution is improved, but it appears that it has diffused away. This fact is consistent with the introduction of diffusive terms in the truncation error which are comparable in magnitude with the diffusivity of the continuous model.

If higher-order asymmetric formulae are used, as the four-points upwind (4PU) defined by

$$\frac{\mathrm{d}T}{\mathrm{d}x} = \frac{T_{j+1} - T_{j-1}}{2h} - q\frac{T_{j+1} - 3T_j + 3T_{j-1} - T_{j-2}}{3h} + O(h^2), \tag{5.1}$$

where q is a free parameter, the accuracy is improved substantially. In Figure 7 we present a numerical experiment obtained with such a method on a uniform grid for $k = 10^{-3}$, $h = 10^{-1}$ and $q = \frac{1}{2}$.

If we compare the profile in Figure 7 with the one obtained in Figure 8 with the same method, but using $q = 0$ – that is central differences on a uniform grid – we observe that 4PU has greatly improved the numerical solution.

Let us return now to the subject of nonuniform grids. If we compare the profile in Figure 7 – obtained with 4PU on a uniform grid with the profile in Figure 4, obtained with Method A on a nonuniform grid – we conclude that method A introduces no dissipation, nor dispersion, in the numerical solution and that 4PU still exhibits some spurious oscillations, and some dissipation.

These experiments suggest that the use of a lower-order discretization formula on a nonuniform grid (method A) produces much better results than a higher-order formula on uniform grids (4PU method). This assertion could, however, invite the following question: can we

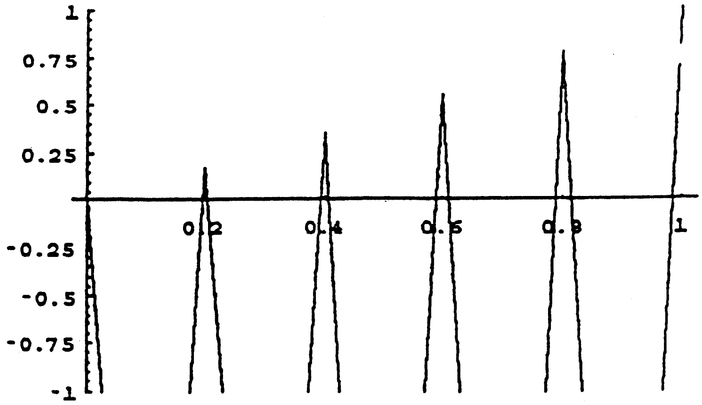

Figure 8. Numerical solution of (2.1) with central differences on a uniform grid, with $h = 10^{-1}$, and $k = 10^{-3}$.

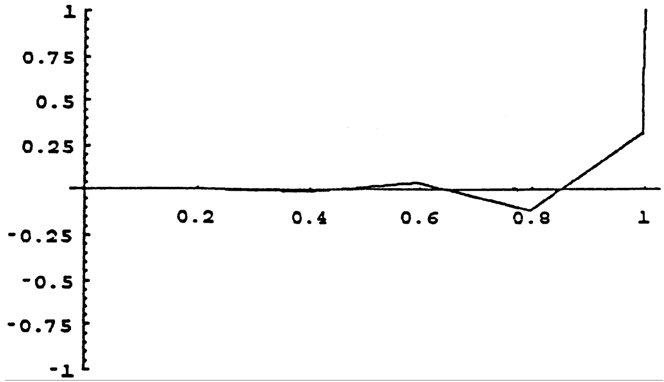

Figure 9. Numerical solution of (2.1) with 4PU on a nonuniform grid, with $N = 10, I = 5, k = 10^{-3}$.

improve the accuracy of the numerical solution when a higher-order formula such as 4PU is used on a nonuniform grid? To answer this question, we have deduced a 4PU method on a nonuniform grid, obtaining

$$\frac{\mathrm{d}T}{\mathrm{d}x} = \frac{T_{j+1} - T_{j-1}}{h_j + h_{j+1}} - \overline{q}_j \frac{\dfrac{h_j T_{j+1} + h_{j+1} T_{j-1} - (h_j + h_{j+1}) T_j}{h_j h_{j+1} h_{j+1/2}} - \dfrac{h_{j-1} T_j + h_j T_{j-2} - (h_j + h_{j-1}) T_{j-1}}{h_j h_{j-1} h_{j-1/2}}}{h_{j+1/2}}. \tag{5.2}$$

We selected q_j in order to cancel the second-order dispersion term on the Modified Equation [3]. We obtain

$$\overline{q}_j = \frac{-(h_{j+1}^3 + h_j^3) + 2k(h_{j+1}^3 - h_j^2)}{4(h_{j-1} + h_j + h_{j+1})}. \tag{5.3}$$

We note that in the case of a uniform grid, $h_j = h$, and we have $\overline{q}_j = \frac{-1}{6} h^2$, which is a value that agrees with the one presented in [3]. Discretizing $\mathrm{d}T/\mathrm{d}x$ with (5.2), (5.3) and $\mathrm{d}^2 T/\mathrm{d}x^2$ with (2.8), we obtain a nonuniform version of 4PU. In Figure 9 we present a numerical experiment obtained with this method for $N = 10, I = 5$ and $k = 10^{-3}$. This numerical experiment, and others that have been carried out, suggest that our question must be answered negatively: the use of a higher-order formula (such as 4PU) on a nonuniform grid hardly improves the result obtained with the same formula on a uniform grid.

Finally we conclude that the most accurate numerical simulations of the two-point boundary-value problem (2.1) – without practically any numerical dispersion nor dissipation – have been obtained with centered finite differences on nonuniform grids. These simulations are much more accurate than those obtained with a higher-order difference formula, such as 4PU, defined on a uniform or a nonuniform grid.

Acknowledgements

The authors thank the anonymous referees whose comments greatly improved the paper.

References

1. A.E.P. Veldman, K. Rinzema, Playing with nonuniform grids. *Journal of Engineering Mathematics* 26 (1992) 119–130.
2. T.A. Manteufel and A.B. White Jr, The numerical solution of second-order boundary-value problems on nonuniform meshes. *math of Comps* 47 (1986) 511–535.
3. C.A.J. Fletcher, *Computational Techniques for Fluid Dynamics 1.* Berlin: Springer-Verlag, (1991).