ELSEVIER

Applied Numerical Mathematics 31 (1999) 463–486

APPLIED NUMERICAL MATHEMATICS

www.elsevier.nl/locate/apnum

Stability of method of characteristics

Mariusz Ziółko *

Department of Electronics AGH, al. Mickiewicza 30, 30-059 Kraków, Poland

Abstract

The Lyapunov functional method is used to verify the stability of a hyperbolic system of two first order partial differential equations. Initial and boundary conditions are assumed. In the case considered Hurwitz type stability occurs simultaneously with Schur type stability. The method of characteristics is used to approximate a continuous system. It is shown that the stability conditions for discrete approximation result from the stability conditions for the continuous system. It is easier to verify the stability of a continuous system than the stability of its discrete approximation. This observation leads to the conclusion that the stability of a continuous system ought to be considered in order to insure convergence of the discrete approximation to the solution of the original problem. A new method for the experimental choice of grid density is proposed. A numerical example is presented in the last part of the paper.

Keywords: Hyperbolic equation; Wave equation; Method of characteristics; Lyapunov method; Hurwitz stability; Schur stability; Grid density

1. Continuous hyperbolic systems

1.1. Hyperbolic systems

Let us consider a system of two linear hyperbolic partial differential equations of the first order

$$\frac{\partial y}{\partial t} + S\frac{\partial y}{\partial x} = Ay, \tag{1}$$

where $S, A \in \mathbb{R}^{2\times 2}$ are constant matrices. Due to the hyperbolicity of (1), all the eigenvalues of the matrix S are real. The wave propagation in a homogeneous one dimensional medium may be described by such a set of differential equations. The solution of (1) has a characteristic property: disturbances propagate at finite speeds. These propagation speeds are equal to the eigenvalues of the matrix S. The positive eigenvalue is a velocity of progressive propagation and the negative one is a velocity of return propagation. Assuming that the absolute values of both speeds are equal, the matrix S can be taken to

* E-mail: ziolko@uci.agh.edu.pl.

0168-9274/99/$20.00
PII: S0168-9274(99)00003-3

the diagonal form $S = \mathrm{diag}\{-s, s\}$. Hyperbolic system (1) has been examined in detail by a number of authors. Its fundamental properties can be found in the classic book [1] by Courant and Hilbert.

C^1 solution of Eq. (1) is uniquely determined inside

$$\Psi = \left\{(t,x)\colon\ 0 \leqslant t \leqslant \frac{1}{2s},\ st \leqslant x \leqslant 1 - st\right\} \tag{2}$$

when C^1 continuous initial values are specified:

$$y(0,x) = y_0(x) \quad \text{for } 0 \leqslant x \leqslant 1. \tag{3}$$

The initial value problem (1), (3) is usually called the Cauchy problem.

Denote by

$$y(t,x) = \begin{bmatrix} y^-(t,x) \\ y^+(t,x) \end{bmatrix} \tag{4}$$

a solution of Eq. (1). It is uniquely determined for

$$\Psi = \{(t,x)\colon\ t \geqslant 0,\ 0 \leqslant x \leqslant 1\} \tag{5}$$

when together with the initial condition (3) the boundary condition

$$\begin{bmatrix} y^-(t,1) \\ y^+(t,0) \end{bmatrix} = B \begin{bmatrix} y^-(t,0) \\ y^+(t,1) \end{bmatrix} \quad \text{for } t \geqslant 0 \tag{6}$$

is imposed. The matrix

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{7}$$

consists of four elements. Two of them, b_{21} and b_{12}, represent boundary reflections at $x = 0$ and $x = 1$, respectively. The elements b_{11} and b_{22} describe respectively a feedback from the boundary $x = 0$ to the boundary $x = 1$ and from the boundary $x = 1$ to the boundary $x = 0$. The initial-boundary value problem (1), (3), (6) is sometimes called the mixed problem.

C^1 regularity, needed for further considerations, is assumed also at the points where the initial condition meets with the boundary condition. This implies the following conditions:

$$\begin{bmatrix} y_0^-(1) \\ y_0^+(0) \end{bmatrix} = B \begin{bmatrix} y_0^-(0) \\ y_0^+(1) \end{bmatrix} \tag{8}$$

and

$$\begin{bmatrix} s\dfrac{\mathrm{d}y_0^-}{\mathrm{d}x}(1) + a_{11}y_0^-(1) + a_{12}y_0^+(1) \\ -s\dfrac{\mathrm{d}y_0^+}{\mathrm{d}x}(0) + a_{21}y_0^-(0) + a_{22}y_0^+(0) \end{bmatrix} = B \begin{bmatrix} s\dfrac{\mathrm{d}y_0^-}{\mathrm{d}x}(0) + a_{11}y_0^-(0) + a_{12}y_0^+(0) \\ -s\dfrac{\mathrm{d}y_0^+}{\mathrm{d}x}(1) + a_{21}y_0^-(1) + a_{22}y_0^+(1) \end{bmatrix}, \tag{9}$$

where $a_{11}, a_{12}, a_{21}, a_{22}$ are elements of matrix A.

1.2. Stability of continuous hyperbolic systems

Definition 1. The dynamic homogeneous system (1), (6) is *asymptotically stable* if

$$E(t) = \|y(t,\cdot)\|^2 \to 0 \quad \text{as } t \to \infty \tag{10}$$

for all $y_0(x)$.

Property (10) can be verified when a norm $\|\cdot\|$ is found such that the derivative $\mathrm{d}E/\mathrm{d}t$ exists and its sign is constant for each $\|y(t,\cdot)\| \neq 0$. If $\mathrm{d}E/\mathrm{d}t < 0$ for all nonzero y, then the system (1), (6) is asymptotically stable. Otherwise, if $\mathrm{d}E/\mathrm{d}t > 0$ for at least one y, the considered system is unstable. Thus we obtain sufficient conditions for the stability or instability of system (1), (6). Combining these two conditions, it is sometimes possible to obtain not only sufficient but also necessary conditions for asymptotic stability. The main difficulty is in proving the sign definiteness of the time derivative with respect to the chosen norm (10). The presented method is quite general and can be applied in other cases, for example of more than two dependent variables (see [9]).

The energy of system (1), (6) can be defined [3] as the square of $L^2_G(0,1)$ norm

$$E(t) = \|y(t,\cdot)\|^2_{L^2_G} = \int_0^1 y^{\mathrm{T}}(t,x)Gy(t,x)\,\mathrm{d}x, \tag{11}$$

where $G \in \mathbb{R}^{2\times 2}$ is a constant positive definite matrix ($G = G^{\mathrm{T}} > 0$). For the initial-boundary value problem (1), (3), (6) the first derivative of energy functional (11) is given by

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \int_0^1 y^{\mathrm{T}}\big(A^{\mathrm{T}}G + GA\big)y\,\mathrm{d}x - y^{\mathrm{T}}SGy\big|_0^1 + \int_0^1 y^{\mathrm{T}}(SG - GS)\frac{\partial y}{\partial x}\,\mathrm{d}x \tag{12}$$

under the assumption that y is a C^1 vector function.

To establish the sign of $\mathrm{d}E/\mathrm{d}t$ we need an additional assumption. If we take the diagonal matrix

$$G = \begin{bmatrix} 1 & \\ & g \end{bmatrix} \tag{13}$$

then the matrices S and G commute, and the bilinear component in formula (12) vanishes. Due to this restriction on matrix G we usually obtain only sufficient conditions for asymptotic stability. When matrix G is diagonal, then

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}E_i}{\mathrm{d}t} + \frac{\mathrm{d}E_b}{\mathrm{d}t}, \tag{14}$$

where

$$\frac{\mathrm{d}E_i}{\mathrm{d}t} = \int_0^1 y^{\mathrm{T}} \begin{bmatrix} 2a_{11} & a_{12} + ga_{21} \\ a_{12} + ga_{21} & 2ga_{22} \end{bmatrix} y\,\mathrm{d}x, \tag{15}$$

$$\frac{\mathrm{d}E_b}{\mathrm{d}t} = s \begin{bmatrix} y^-(t,0) \\ y^+(t,1) \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} b_{11}^2 + gb_{21}^2 - 1 & b_{11}b_{12} + gb_{21}b_{22} \\ b_{11}b_{12} + gb_{21}b_{22} & b_{12}^2 + gb_{22}^2 - g \end{bmatrix} \begin{bmatrix} y^-(t,0) \\ y^+(t,1) \end{bmatrix}. \tag{16}$$

The first derivative of the energy functional (11) has been split into two parts: (15) and (16). The first part depends on matrix A and the second on matrix B. Matrix A is a coefficient of differential equation (1) describing the dynamics of the system inside Ψ defined by (5). On the other hand, matrix B is defined by boundary condition (6) and describes properties of the variable y at the boundaries of Ψ. These observations motivate the following definitions.

Definition 2. The initial-boundary value problem (1), (3), (6) is *asymptotically interior stable* if the inequality $\mathrm{d}E_i/\mathrm{d}t < 0$ holds for all $t \geqslant 0$ and for all non-zero initial conditions (3).

Definition 3. The initial-boundary value problem (1), (3), (6) is *asymptotically boundary stable* if $\mathrm{d}E_b/\mathrm{d}t < 0$ for all $t \geqslant 0$ and for y not equal to zero at both boundaries simultaneously.

Definition 4. A matrix (polynomial) is Hurwitz stable if all its eigenvalues (zeros) have negative real parts.

Definition 5. A matrix (polynomial) is Schur stable if all its eigenvalues (zeros) have magnitudes less than 1.

Lemma 1. *Matrix $A \in \mathbb{R}^{2\times 2}$ is Hurwitz stable if and only if* $\operatorname{tr} A < 0$ *and* $\det A > 0$.

Proof. The eigenvalues of matrix $A \in \mathbb{R}^{2\times 2}$ have negative real parts if and only if the characteristic polynomial

$$\det[A - \lambda I] = \lambda^2 - \lambda \operatorname{tr} A + \det A \tag{17}$$

has positive coefficients. □

Theorem 1. *If* $\det A > 0$ *and* $a_{11} < 0$ $(a_{22} < 0)$ *then the initial-boundary value problem* (1), (3), (6) *is asymptotically interior stable, and parameter g in* (15) *satisfies the inequalities*

$$\frac{(\sqrt{\det A} - \sqrt{a_{11}a_{22}}\,)^2}{a_{21}^2} < g < \frac{(\sqrt{\det A} + \sqrt{a_{11}a_{22}}\,)^2}{a_{21}^2} \quad \textit{if } a_{21} \neq 0 \tag{18}$$

or

$$g > \frac{a_{12}^2}{4a_{11}a_{22}} \quad \textit{if } a_{21} = 0. \tag{19}$$

Proof. The matrix

$$\begin{bmatrix} 2a_{11} & a_{12} + ga_{21} \\ a_{12} + ga_{21} & 2ga_{22} \end{bmatrix} \tag{20}$$

is negative definite if its minors of the first order are negative (i.e., $a_{11} < 0$ and $a_{22} < 0$) and the determinant is positive, i.e.,

$$-g^2a_{21}^2 + 2g(\det A + a_{11}a_{22}) - a_{12}^2 > 0. \tag{21}$$

For the case $a_{21} \neq 0$ inequality (21) is satisfied if g fulfills condition (18). If $a_{21} = 0$ inequality (21) is equivalent to (19). □

Lemma 2. *The matrix* $B \in \mathbb{R}^{2\times 2}$ *is Schur stable if and only if*

$$|\mathrm{tr} B| - 1 < \det B < 1. \tag{22}$$

Proof. The transformation $\lambda = (z+1)/(z-1)$ puts the characteristic polynomial $\det[B - \lambda I]$ into the form

$$(1 - \mathrm{tr} B + \det B)z^2 + 2(1 - \det B)z + 1 + \mathrm{tr} B + \det B. \tag{23}$$

Polynomial (23) is Hurwitz stable if and only if its coefficients are positive. □

Theorem 2. *If*

$$|\mathrm{tr} B| < 1 + \det B, \tag{24}$$

$$b_{11}^2 < 1, \tag{25}$$

$$b_{12}^2 b_{21}^2 < (1 - b_{11}^2)(1 - b_{22}^2), \tag{26}$$

$$\det B < 1 - |b_{11} - b_{22}| \quad \text{if } b_{21} \neq 0, \tag{27}$$

then the initial-boundary value problem (1), (3), (6) *is asymptotically boundary stable and parameter* g *satisfies the inequalities*

$$0 \leqslant \max\left\{\frac{b_{12}^2}{1 - b_{22}^2}, \frac{\alpha - \sqrt{\Delta}}{2b_{21}^2}\right\} < g < \min\left\{\frac{1 - b_{11}^2}{b_{21}^2}, \frac{\alpha + \sqrt{\Delta}}{2b_{21}^2}\right\} \tag{28}$$

if $b_{21} \neq 0$ *or*

$$g > \frac{b_{12}^2}{(1 - b_{11}^2)(1 - b_{22}^2)} \quad \text{if } b_{21} = 0, \tag{29}$$

where

$$\alpha = \det{}^2 B + 1 - b_{11}^2 - b_{22}^2, \tag{30}$$

$$\Delta = \alpha^2 - 4b_{12}^2 b_{21}^2. \tag{31}$$

Proof. First we shall prove the theorem for the case $b_{21} \neq 0$. From inequalities (26) and (25) it follows that there exists $g > 0$ such that

$$\frac{b_{12}^2}{1 - b_{22}^2} < g < \frac{1 - b_{11}^2}{b_{21}^2}. \tag{32}$$

This implies that the elements on the diagonal in (16) are negative. To complete this part of the proof it remains to be shown that the determinant of matrix in (16) is positive. After simple calculations we obtain

$$g^2 b_{21}^2 - g\alpha + b_{12}^2 < 0. \tag{33}$$

This inequality is satisfied if the discriminant of the polynomial in (33) is positive

$$\Delta = \left[(b_{11} - b_{22})^2 - (1 - \det B)^2\right]\left[\mathrm{tr}^2 B - (1 + \det B)^2\right] > 0. \tag{34}$$

The discriminant Δ is a product of two negative elements. The first of them is negative due to (27) and the second due to (24). Inequality (33) holds whenever

$$\frac{\alpha - \sqrt{\Delta}}{2b_{21}^2} < g < \frac{\alpha + \sqrt{\Delta}}{2b_{21}^2}. \tag{35}$$

From (31) it is clear that both sides of (35) are not negative. To end this part of the proof we shall notice that there exists $g > 0$ such that both inequalities (32) and (35) are satisfied simultaneously. For the case $b_{21} = 0$ condition (33) takes the form

$$g\left(1 - b_{11}^2\right)\left(1 - b_{22}^2\right) - b_{12}^2 > 0 \tag{36}$$

which by assumption (26) is equivalent to (29). □

The Lyapunov functional for the Cauchy problem has the form

$$E(t) = \left\| y(t,\cdot)\right\|_{L_G^2}^2 = \int_{st}^{1-st} y^{\mathrm{T}}(t,x) G y(t,x)\,\mathrm{d}x \tag{37}$$

and its first derivative

$$\begin{aligned}\frac{\mathrm{d}E}{\mathrm{d}t} = {}& y^{\mathrm{T}}(t,1-st)\begin{bmatrix} 0 & 0 \\ 0 & -2sg \end{bmatrix} y(t,1-st) + y^{\mathrm{T}}(t,st)\begin{bmatrix} -2s & 0 \\ 0 & 0 \end{bmatrix} y(t,st) \\ &+ \int_{st}^{1-st} y^{\mathrm{T}}(t,x)\left(A^{\mathrm{T}}G + GA\right) y(t,x)\,\mathrm{d}x. \end{aligned}\tag{38}$$

Definition 6. The Cauchy problem (1), (3) is *asymptotically stable* if there exists $g > 0$ such that the inequality $\mathrm{d}E/\mathrm{d}t < 0$ holds for all $0 \leqslant t < 0.5/s$ and all nonzero initial conditions (3).

Theorem 3. *The asymptotic stability of the Cauchy problem* (1), (3) *is equivalent to the asymptotic interior stability of the initial-boundary value problem* (1), (3), (6).

Proof. The first two quadratic forms in (38) are negative semi-definite. Negative definiteness of the quadratic form $A^{\mathrm{T}}G + GA$ is a sufficient condition for both the Cauchy problem (1), (3) and the mixed problem (1), (3), (6). To prove the necessity of this condition let us notice that $\mathrm{d}E/\mathrm{d}t < 0$ for *all* initial conditions implies the negative definiteness of the quadratic form $A^{\mathrm{T}}G + GA$. Indeed, this implication must hold also for initial conditions (3) which makes the first two quadratic forms in (38) equal to zero. Such initial conditions exist and can be determined when $y^-(t, 1-st) \neq 0$, $y^+(t,st) \neq 0$ and $y^+(t, 1-st) = y^-(t,st) = 0$ along the characteristics defined by $-s$ and s. □

2. Discrete hyperbolic systems

2.1. Method of characteristics

The method of characteristics is the main finite difference method for computing the numerical solution of partial differential equations of hyperbolic type. Its efficiency results from the proper grid points allocation which reflects the "wave nature" of hyperbolic equations. The algorithm is simple even for nonstationary or quasi-linear equations (see [1]). This simplicity reduces the time of computations. Along the characteristic curves

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -s, \qquad \frac{\mathrm{d}x}{\mathrm{d}t} = s \tag{39}$$

the partial differential equations

$$\begin{aligned} \frac{\partial y^-}{\partial t} - s\frac{\partial y^-}{\partial x} &= a_{11}y^- + a_{12}y^+, \\ \frac{\partial y^+}{\partial t} + s\frac{\partial y^+}{\partial x} &= a_{21}y^- + a_{22}y^+ \end{aligned} \tag{40}$$

are transformed into the set of the ordinary differential equations

$$\begin{aligned} \frac{\mathrm{d}y^-}{\mathrm{d}t} &= a_{11}y^- + a_{12}y^+, \\ \frac{\mathrm{d}y^+}{\mathrm{d}t} &= a_{21}y^- + a_{22}y^+. \end{aligned} \tag{41}$$

At this stage, a simple finite difference method is used to solve numerically Eqs. (41). In the case of two-level methods, the difference equations can be written in the matrix notation

$$y_{j+1,i} = \Gamma y_{j,i} + \Omega y_{j,i+1}, \tag{42}$$

where $y_{j,i} = [y^-_{j,i}\ y^+_{j,i}]^{\mathrm{T}} \in \mathbb{R}^2$. The numbering of nodes is explained in Fig. 1(a). The first index denotes the time and the second the space discrete variable. For the Euler method,

$$\Gamma = \begin{bmatrix} 0 & 0 \\ \Delta t a_{21} & 1+\Delta t a_{22} \end{bmatrix}, \qquad \Omega = \begin{bmatrix} 1+\Delta t a_{11} & \Delta t a_{12} \\ 0 & 0 \end{bmatrix}. \tag{43}$$

Matrices Γ and Ω are singular, each matrix has one eigenvalue equal to zero.

2.2. *Stability of discrete Cauchy problem*

The method of characteristics enables one to obtain a discrete approximation of the solution to problem (1), (3). Mesh length is determined by the size of space discretization Δx and has to be chosen so that $1/\Delta x$ is an integer. Then the discrete domain

$$\begin{aligned} \Psi_d = \big\{(t_j, x_i) \in \Psi\colon\ & t_j = 0.5(j-1)\Delta x/s;\ x_i = \Delta x[i-1+0.5(j-1)]; \\ & j = 1,2,\ldots,1+1/\Delta x;\ i = 1,2,\ldots,2-j+1/\Delta x\big\}, \end{aligned} \tag{44}$$

where Ψ is defined by (2). The discrete values in the first time level are determined by initial condition (3)

$$y_{1,i} = y_0(x_i), \tag{45}$$

where $i = 1,2,\ldots,1+1/\Delta x$ and $x_i = (i-1)\Delta x$. For $j > 1$ the recursive formula (42) is applied and the iterative procedure can be written in the matrix notation

$$y_{j+1} = \Phi_c\, y_j, \tag{46}$$

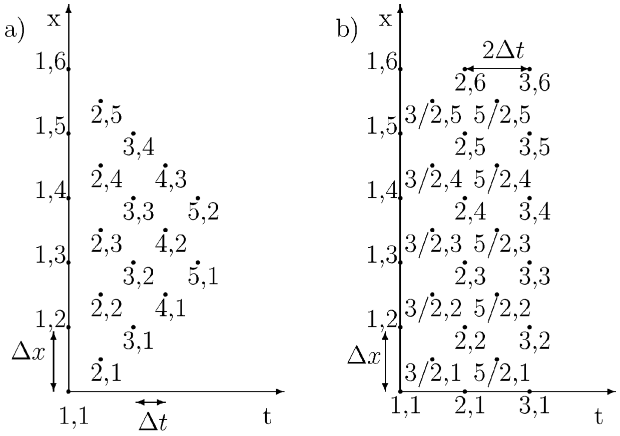

Fig. 1. Numbering of nodes for the discrete: (a) Cauchy problem, (b) initial-boundary value problem.

where

$$y_j = \begin{bmatrix} y_{j,1} \\ \vdots \\ y_{j,I} \end{bmatrix}, \quad I = 2 - j + 1/\Delta x, \qquad y_j \in \mathbb{R}^{2I}, \tag{47}$$

and rectangular matrix Φ_c consists of matrices Γ and Ω:

$$\Phi_c = \begin{bmatrix} \Gamma & \Omega & & & \\ & \Gamma & \Omega & & \\ & & \cdots & \cdots & \\ & & & \Gamma & \Omega \end{bmatrix}, \tag{48}$$

where $\dim \Phi_c = 2(1 - j + 1/\Delta x) \times 2(2 - j + 1/\Delta x)$. Matrices Γ, Ω are defined by (43) according to the Euler method chosen for solving the ordinary differential equations (41) along the characteristics (39). Formula (46) constitutes a discrete approximation of the partial differential equation (1). As for the continuous system, it is possible to define the energy functional

$$E_j = \|y_j\|^2_{l^2_{G_j}} = y_j^{\mathrm{T}} G_j y_j \tag{49}$$

for discrete system (46) where the dimensions of diagonal matrices G_j depend on the number of time step j

$$\dim G_j = 2(2 - j + 1/\Delta x) \times 2(2 - j + 1/\Delta x). \tag{50}$$

Matrices G_j can be defined so that the both norms for the continuous-time system and the discrete-time system are consistent in the following sense:

$$\lim_{\Delta x \to 0} \Delta x \|y_j\|^2_{l^2_{G_j}} = \|y(t,\cdot)\|^2_{L^2_G}, \tag{51}$$

where matrices G_j are block diagonal and the elements on diagonal are $G = \mathrm{diag}\{1, g\}$.

Definition 7. The discrete linear system (46) is *asymptotically stable* if there exists such a matrix $G_j = \operatorname{diag}(G, \dots, G) \in \mathbb{R}^{2I \times 2I}$ that $y_{j+1}^{\mathrm{T}} G_{j+1} y_{j+1} - y_j^{\mathrm{T}} G_j y_j < 0$ for all $j = 1, 2, \dots, 1/\Delta x$ and every y_j with $\|y_1\| \neq 0$.

Definition 8. The discrete system (46) is *numerically stable* if there exists such a positive constant τ that (46) is asymptotically stable for all $0 < \Delta t < \tau$.

Proposition 1. *If there exists a sequence of such non-singular matrices H_j that* $\dim H_j = 2(2 - j + 1/\Delta x) \times 2(2 - j + 1/\Delta x)$ *and*

$$\left\| H_{j+1} \Phi_c H_j^{-1} \right\| < 1 \tag{52}$$

for all $1 \leqslant j \leqslant 1 + 1/\Delta x$, then the discrete system $y_{j+1} = \Phi_c \, y_j$ is asymptotically stable.

Proof. By the definition of matrix spectral norm and assumption (52) it follows that

$$\frac{z_j^{\mathrm{T}} (H_j^{-1})^{\mathrm{T}} \Phi_c^{\mathrm{T}} H_{j+1}^{\mathrm{T}} H_{j+1} \Phi_c H_j^{-1} z_j}{z_j^{\mathrm{T}} z_j} < 1 \tag{53}$$

for all $z_j \neq 0$. Substituting

$$z_j = H_j y_j, \qquad G_{j+1} = H_{j+1}^{\mathrm{T}} H_{j+1}, \qquad G_j = H_j^{\mathrm{T}} H_j \tag{54}$$

into (53) we obtain

$$y_j^{\mathrm{T}} \Phi_c^{\mathrm{T}} G_{j+1} \Phi_c y_j < y_j^{\mathrm{T}} G_j y_j \tag{55}$$

and finally taking into account (46)

$$\|y_{j+1}\|^2_{l^2_{G_{j+1}}} < \|y_j\|^2_{l^2_{G_j}}. \tag{56}$$

□

Theorem 4. *If the continuous Cauchy problem* (1), (3) *is asymptotically stable, i.e.,*

$$a_{11} < 0, \qquad a_{22} < 0, \qquad \det A > 0, \tag{57}$$

then discrete system (46)*, resulting from the Euler method applied to the method of characteristics, is numerically stable.*

Proof. To use Proposition 1 let us take $H_{j+1} = G_{j+1}^{1/2}$ and $H_j = G_j^{1/2}$, where diagonal matrices G_j and G_{j+1} are defined by (49)–(51) and $1/2$ denotes a power exponent. The norm of matrix $G_{j+1}^{1/2} \Phi_c G_j^{-1/2}$ is less than 1 if all the eigenvalues of the matrix

$$\Theta_c = G_{j+1}^{1/2} \Phi_c G_j^{-1} \Phi_c^{\mathrm{T}} G_{j+1}^{1/2} \tag{58}$$

are in [0, 1), where matrix Φ_c is defined by (43) and (48). To verify this let us consider the determinant

$$\varXi_{2n} = \det(\Theta_c - \lambda I) = \det \begin{bmatrix} a & 0 & 0 & b & 0 & 0 & 0 & 0 & \dots & & & & \\ 0 & c & 0 & 0 & 0 & 0 & 0 & 0 & \dots & & & & \\ 0 & 0 & a & 0 & 0 & b & 0 & 0 & \dots & & & & \\ b & 0 & 0 & c & 0 & 0 & 0 & 0 & \dots & & & & \\ 0 & 0 & 0 & 0 & a & 0 & 0 & b & \dots & & & & \\ 0 & 0 & b & 0 & 0 & c & 0 & 0 & \dots & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & a & 0 & \dots & & & & \\ 0 & 0 & 0 & 0 & b & 0 & 0 & c & \dots & & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ & & & & & & & & \dots & a & 0 & 0 & b \\ & & & & & & & & \dots & 0 & c & 0 & 0 \\ & & & & & & & & \dots & 0 & 0 & a & 0 \\ & & & & & & & & \dots & b & 0 & 0 & c \end{bmatrix}, \tag{59}$$

where

$$\begin{aligned} a &= (1 + \Delta t a_{11})^2 + a_{12}^2 \Delta t^2 / g - \lambda, \\ b &= (1 + \Delta t a_{11}) \Delta t a_{21} \sqrt{g} + (1 + \Delta t a_{22}) \Delta t a_{12} / \sqrt{g}, \\ c &= \Delta t^2 a_{21}^2 g + (1 + \Delta t a_{22})^2 - \lambda, \\ n &= (1/\Delta x - j + 1), \qquad \dim \Theta_c = 2n \times 2n. \end{aligned} \tag{60}$$

We will denote as $\varXi_{2n-k}$ a minor which results from determinant $\varXi_{2n}$ by deleting the first k rows and k columns. $\varPi_{2n}$ will denote a minor obtained from determinant $\varXi_{2n}$ by deleting the sixth and the first four columns and rows. With this notation we derived the set of equations

$$\begin{aligned} \varXi_{2n} &= a \varXi_{2n-1} - a b^2 c \varXi_{2n-4} + b^4 c \varPi_{2n}, \\ \varXi_{2n-1} &= c \varXi_{2n-2}, \\ \varXi_{2n-2} &= a \varXi_{2n-3} - b^2 c \varPi_{2n}, \\ \varXi_{2n-3} &= c \varXi_{2n-4}. \end{aligned} \tag{61}$$

By eliminating the minors of odd range and minor $\varPi_{2n}$ we obtain

$$\varXi_{2n} + (b^2 - ac) \varXi_{2n-2} = 0. \tag{62}$$

The characteristic equation for (62) has the form

$$r^2 + (b^2 - ac) = 0 \tag{63}$$

and the general solution of (62) is given by

$$\Xi_{2n} = \alpha r^{2n} = \alpha\left(ac - b^2\right)^n. \tag{64}$$

The constant α can be calculated from the value of determinant $\Xi_2 = ac$. Thus we obtain

$$\Xi_{2n} = ac\left(ac - b^2\right)^{n-1}. \tag{65}$$

Zeros of determinant (65) give two eigenvalues of Θ_c:

$$\lambda_1 = (1 + \Delta t a_{11})^2 + a_{12}^2 \Delta t^2/g, \qquad \lambda_2 = (1 + \Delta t a_{22})^2 + a_{21}^2 \Delta t^2 g \tag{66}$$

which are less than one if respectively

$$\Delta t < \frac{-2a_{11}}{a_{11}^2 + a_{12}^2/g}, \qquad \Delta t < \frac{-2a_{22}}{a_{22}^2 + a_{21}^2 g}. \tag{67}$$

By the assumption $a_{11} < 0$ we obtain $\lambda_1 < 1$ for sufficiently small Δt. Similarly from the inequality $a_{22} < 0$ it follows that $\lambda_2 < 1$. The other eigenvalues of matrix Θ_c are solutions of the equation

$$\begin{aligned} ac - b^2 = {} & \lambda^2 - \lambda\left[\Delta t^2\left(a_{11}^2 + a_{22}^2 + a_{12}^2/g + a_{21}^2 g\right) + 2\Delta t \operatorname{tr} A + 2\right] \\ & + \left[\Delta t^2 \det A + \Delta t \operatorname{tr} A + 1\right]^2 = 0. \end{aligned} \tag{68}$$

The zeros of polynomial (68) are real and non-negative because the symmetric matrix Θ_c is non-negative definite. To verify that they belong to segment $[0, 1)$, let us shift the co-ordinate system by substituting $\lambda = \varepsilon + 1$. In this way we obtain the equation

$$\begin{aligned} 0 = {} & \varepsilon^2 - \varepsilon\left[\Delta t^2\left(a_{11}^2 + a_{22}^2 + a_{12}^2/g + a_{21}^2 g\right) + 2\Delta t \operatorname{tr} A\right] \\ & + \Delta t^2\left(\Delta t^2 \det^2 A + 2\Delta t \operatorname{tr} A \det A + 2 \det A + 2a_{11}a_{22} - a_{12}^2/g - a_{21}^2 g\right). \end{aligned} \tag{69}$$

Both zeros of (69) belong to segment $[-1, 0)$ if the polynomial coefficients are positive. This means that we obtain two conditions:

$$\Delta t\left(a_{11}^2 + a_{22}^2 + a_{12}^2/g + a_{21}^2 g\right) + 2 \operatorname{tr} A < 0, \tag{70}$$

$$\Delta t^2 \det^2 A + 2\Delta t \operatorname{tr} A \det A + 2 \det A + 2a_{11}a_{22} - a_{12}^2/g - g a_{21}^2 > 0. \tag{71}$$

From (67) it follows that inequality (70) is satisfied. Otherwise the two inequalities must be satisfied simultaneously:

$$\begin{aligned} \frac{-2a_{11} - 2a_{22}}{a_{11}^2 + a_{22}^2 + a_{12}^2/g + a_{21}^2 g} &< \frac{-2a_{11}}{a_{11}^2 + a_{12}^2/g}, \\ \frac{-2a_{11} - 2a_{22}}{a_{11}^2 + a_{22}^2 + a_{12}^2/g + a_{21}^2 g} &< \frac{-2a_{22}}{a_{22}^2 + a_{21}^2 g}. \end{aligned} \tag{72}$$

If an appropriate transformation is taken, we notice that (72) consists of contradictory inequalities. It remains to verify inequality (71). The polynomial in (71) has two real roots

$$\Delta t_1 = \frac{-\operatorname{tr} A - \sqrt{(a_{11} - a_{22})^2 + (a_{12}/\sqrt{g} + a_{21}\sqrt{g})^2}}{\det A}, \tag{73}$$

$$\Delta t_2 = \frac{-\operatorname{tr} A + \sqrt{(a_{11} - a_{22})^2 + (a_{12}/\sqrt{g} + a_{21}\sqrt{g})^2}}{\det A}. \tag{74}$$

From $\mathrm{tr}A < 0$ and $\det A > 0$ we obtain immediately $\Delta t_2 > 0$. Thus it remains to be verified that $\Delta t_1 > 0$. Δt_1 depends on g and takes its maximum value

$$\Delta t_1 = \frac{-\mathrm{tr}A - \sqrt{(a_{11} - a_{22})^2 + 2(|a_{12}a_{21}| + a_{12}a_{21})}}{\det A} \tag{75}$$

when

$$g = \left|\frac{a_{12}}{a_{21}}\right|. \tag{76}$$

If $a_{12}a_{21} < 0$, the numerator in (75) is positive since $a_{11} < 0$ and $a_{22} < 0$. In the opposite case $a_{12}a_{21} > 0$, the numerator in (75) is equal to

$$-\mathrm{tr}A - \sqrt{\mathrm{tr}^2 A - 4\det A} \tag{77}$$

and from (57) it follows that it is positive, too. To complete the proof we have to consider the special cases when parameter g cannot be found from (76). For $a_{12} = 0$ we obtain

$$\Delta t_1 = \frac{-\mathrm{tr}A - \sqrt{(a_{11} - a_{22})^2 + a_{21}^2 g}}{a_{11}a_{22}} \tag{78}$$

and condition $\Delta t_1 > 0$ is satisfied if

$$g < \frac{4a_{11}a_{22}}{a_{21}^2}. \tag{79}$$

For the case $a_{21} = 0$ we obtain

$$\Delta t_1 = \frac{-\mathrm{tr}A - \sqrt{(a_{11} - a_{22})^2 + a_{12}^2/g}}{a_{11}a_{22}} \tag{80}$$

and condition $\Delta t_1 > 0$ is satisfied if

$$g > \frac{a_{12}^2}{4a_{11}a_{22}}. \tag{81}$$

For the case $a_{12} = a_{21} = 0$ we obtain

$$\Delta t_1 = -2/a_{11}, \qquad \Delta t_2 = -2/a_{22}. \tag{82}$$

□

From Theorem 4 we deduce that there are two sufficient conditions for the asymptotic stability of the discrete system (46): the asymptotic stability of the continuous system (40) and the sufficiently small density of discretization

$$\Delta t < \max_g \min\left\{\frac{-2a_{11}}{a_{11}^2 + a_{12}^2/g}, \frac{-2a_{22}}{a_{22}^2 + a_{21}^2 g}, \frac{-\mathrm{tr}A - \sqrt{(a_{11} - a_{22})^2 + g(a_{12}/g + a_{21})^2}}{\det A}\right\}, \tag{83}$$

where the minimum is taken out of the three elements. The first two elements are always positive for the stable continuous Cauchy problem. The third element is positive if the parameter g is chosen properly. The maximum value of the third element is equal to (75) for (76).

2.3. Stability of discrete initial-boundary value problem

The discrete approximation of initial-boundary value problem (1), (3), (6) is calculated at the nodes (t_j, x_i) which form the domain

$$\Psi_d = \{(t_j, x_i) \in \Psi\colon\ t_j = 2(j-1)\Delta t;\ x_i = (i-1)\Delta x;\ \Delta x = 2s\Delta t;\\ j = 1, 2, \ldots, 1 + sT/\Delta x;\ i = 1, 2, \ldots, 1 + 1/\Delta x\}, \tag{84}$$

where space discretization Δx and final time T are chosen so that $1/\Delta x$ and $sT/\Delta x$ are integers. The grid numbering is presented in Fig. 1(b).

The discrete values of the first time-level are known from initial condition (3):

$$y_{i,1} = y_0(x_i), \tag{85}$$

where $i = 1, 2, \ldots, 1 + 1/\Delta x$ and $x_i = (i-1)\Delta x$. Next, the values of the discrete solution in the intermediate time level, i.e., $j = \frac{3}{2}$, are calculated. The corresponding space co-ordinates for these nodes are $\frac{1}{2}\Delta x,\ \frac{3}{2}\Delta x,\ \frac{5}{2}\Delta x, \ldots, 1 - \frac{1}{2}\Delta x$. The values $y_{3/2,i}$ (where $i = 1, 2, \ldots, 1/\Delta x$) do not depend on boundary condition (6) and are calculated in the same way as for the Cauchy problem. This part of the calculation can be written in the matrix notation

$$y_{j+1/2} = \Phi_c\, y_j, \tag{86}$$

where rectangular matrix Φ_c is defined by (48) and dim $\Phi_c = 2/\Delta x \times 2(1 + 1/\Delta x)$ does not depend on discrete time j.

In the same way, the values for the next time level are calculated (i.e., $j = 2$) for the nodes $x_i = i\Delta x$ where $i = 2, 3, \ldots, 1/\Delta x$. The values $y_{2,1}$ and $y_{2,1+1/\Delta x}$ for the boundary points are calculated from both the boundary conditions (6) and difference equations (42). This leads to the difference equations

$$\begin{cases} y^-_{j,1} - y^-_{j-1/2,3/2} = (a_{11} y^-_{j-1/2,3/2} + a_{12} y^+_{j-1/2,3/2})\Delta t, \\ y^+_{j,1} = b_{21} y^-_{j,1} + b_{22} y^+_{j,m}, \end{cases} \tag{87}$$

$$\begin{cases} y^-_{j,m} = b_{12} y^+_{j,m} + b_{11} y^-_{j,1}, \\ y^+_{j,m} - y^+_{j-1/2,n} = (a_{21} y^-_{j-1/2,n} + a_{22} y^+_{j-1/2,n})\Delta t, \end{cases} \tag{88}$$

where $m = 1 + 1/\Delta x$ and $n = \frac{1}{2} + 1/\Delta x$. From these equations we obtain values at the boundary nodes

$$y_{j,1} = V y_{j-1/2,3/2} + C y_{j-1/2,n}, \qquad y_{j,m} = W y_{j-1/2,n} + D y_{j-1/2,3/2}, \tag{89}$$

where

$$V = \begin{bmatrix} 1 + \Delta t a_{11} & \Delta t a_{12} \\ b_{21}(1 + \Delta t a_{11}) & b_{21}\Delta t a_{12} \end{bmatrix}, \qquad W = \begin{bmatrix} b_{12}\Delta t a_{21} & b_{12}(1 + \Delta t a_{22}) \\ \Delta t a_{21} & 1 + \Delta t a_{22} \end{bmatrix},\\ C = \begin{bmatrix} 0 & 0 \\ \Delta t b_{22} a_{21} & b_{22}(1 + \Delta t a_{22}) \end{bmatrix}, \qquad D = \begin{bmatrix} b_{11}(1 + \Delta t a_{11}) & \Delta t b_{11} a_{12} \\ 0 & 0 \end{bmatrix}. \tag{90}$$

This part of the calculation can be written in the matrix notation

$$y_{j+1} = \Phi_b\, y_{j+1/2}, \tag{91}$$

where

$$\Phi_b = \begin{bmatrix} V & & & & C \\ \Gamma & \Omega & & & \\ & \Gamma & \Omega & & \\ & & \dots & \dots & \\ & & & \Gamma & \Omega \\ D & & & & W \end{bmatrix} \tag{92}$$

and dim $\Phi_b = 2(1 + 1/\Delta x) \times 2/\Delta x$ does not depend on the number of time level. Square matrices Γ and Ω are defined by (43).

Conjecture 1. Discrete system (91) is numerically stable if the continuous initial-boundary value problem (1), (3), (6) is asymptotically interior and boundary stable.

Justification is presented in the Appendix.

The final formula for the approximation of the initial-boundary value problem is the composition of both (86) and (91):

$$y_{j+1} = \Phi y_j, \tag{93}$$

where

$$\Phi = \Phi_b \Phi_c = \begin{bmatrix} V\Gamma & V\Omega & & & & C\Gamma & C\Omega \\ \Gamma\Gamma & \Gamma\Omega + \Omega\Gamma & \Omega\Omega & & & & \\ & \Gamma\Gamma & \Gamma\Omega + \Omega\Gamma & \Omega\Omega & & & \\ & & \dots & \dots & \dots & & \\ & & & \dots & \dots & \dots & \\ & & & & \Gamma\Gamma & \Gamma\Omega + \Omega\Gamma & \Omega\Omega \\ D\Gamma & D\Omega & & & & W\Gamma & W\Omega \end{bmatrix} \tag{94}$$

dim $\Phi = 2(1 + 1/\Delta x) \times 2(1 + 1/\Delta x)$. If (86) and (91) are asymptotically stable then their superposition (93) is also asymptotically stable.

3. Experimental choice of grid density

Let us consider the interior and boundary stable initial-boundary value problem

$$\frac{\partial y}{\partial t} + S\frac{\partial y}{\partial x} = Ay, \tag{95}$$

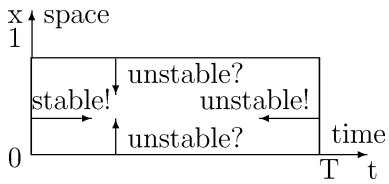

Fig. 2. The stability of the mathematical models depends on the direction of computations.

$$\begin{bmatrix} y^-(t,1) \\ y^+(t,0) \end{bmatrix} = B \begin{bmatrix} y^-(t,0) \\ y^+(t,1) \end{bmatrix}, \tag{96}$$

where

$$A = \overrightarrow{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad B = \overrightarrow{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \qquad S = \overrightarrow{S} = \begin{bmatrix} -s & \\ & s \end{bmatrix}. \tag{97}$$

Now, our goal is to investigate the stability of the system described by Eqs. (95), (96) when the final condition

$$y(T,x) = y_T(x) \tag{98}$$

is given instead of the initial condition (3). In other words, Eq. (95) ought to be integrated backwards. The final-boundary value problem (95), (96), (98) can be transformed to its canonical form. After a simple calculation we obtain new matrix coefficients for differential equation (95) and boundary condition (96):

$$\begin{aligned} &A = \overleftarrow{A} = \begin{bmatrix} -a_{22} & -a_{21} \\ -a_{12} & -a_{11} \end{bmatrix}, \qquad B = \overleftarrow{B} = \frac{1}{\det \overrightarrow{B}} \begin{bmatrix} b_{11} & -b_{21} \\ -b_{12} & b_{22} \end{bmatrix}, \\ &S = \overleftarrow{S} = \overrightarrow{S} = \begin{bmatrix} -s & \\ & s \end{bmatrix}. \end{aligned} \tag{99}$$

Proposition 2. *If system* (95)–(97) *is interior and boundary stable then the corresponding final-boundary value problem* (95), (96), (99) *is interior and boundary unstable.*

Proof. The interior stability of the initial-boundary value problem (95)–(97) implies that $a_{11} < 0$ and $a_{22} < 0$. This makes the elements on the diagonal of the matrix $\overleftarrow{A}$ positive, and is contradictory to the assumptions of Theorem 1. In this way we come to the conclusion that the final-boundary value problem (95), (96) with coefficients (99) is interior unstable. Taking into account that

$$\det \overrightarrow{B} = \frac{1}{\det \overleftarrow{B}}, \qquad \operatorname{tr} \overrightarrow{B} = \frac{\operatorname{tr} \overleftarrow{B}}{\det \overleftarrow{B}} \tag{100}$$

we deduce from the Schur stability of matrix $\overrightarrow{B}$ that the final-boundary value problem is boundary unstable. □

For an interior stable system both eigenvalues $\lambda_i(\overrightarrow{A})$ $(i = 1, 2)$ have negative real parts. The change of time direction increases the real parts of the eigenvalues of matrix $\overleftarrow{A}$,

$$\lambda_i(\overleftarrow{A}) = \lambda_i(\overrightarrow{A}) + |\mathrm{tr}\,\overrightarrow{A}|, \tag{101}$$

and they become positive. For the boundary stable system (95), (96) eigenvalues of matrix $\overrightarrow{B}$ have magnitudes less than 1. The corresponding final-boundary value problem does not preserve this property and the magnitudes of the eigenvalues $\lambda_i(\overleftarrow{B})$ are greater than 1.

Now let us consider the boundary condition

$$y(t, 0) = y_0(t) \quad \text{for } 0 \leqslant t \leqslant T. \tag{102}$$

The boundary value problem (95), (102) is similar to the Cauchy problem (1), (3). The main difference is in the different role of the time and space variables. Hence we obtain new coefficients for the canonical form of Eq. (95):

$$S = S_{\uparrow} = \begin{bmatrix} -1/s & \\ & 1/s \end{bmatrix}, \qquad A = A_{\uparrow} = \frac{1}{s}\begin{bmatrix} -a_{11} & -a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \tag{103}$$

The matrix $A_{\uparrow}$ has two eigenvalues

$$\begin{aligned} \lambda_1(A_{\uparrow}) &= \frac{a_{22} - a_{11} - \sqrt{(a_{22} - a_{11})^2 + 4\det \overrightarrow{A}}}{2s}, \\ \lambda_2(A_{\uparrow}) &= \frac{a_{22} - a_{11} + \sqrt{(a_{22} - a_{11})^2 + 4\det \overrightarrow{A}}}{2s}. \end{aligned} \tag{104}$$

For the interior stable system (95), (97) (i.e., $\det \overrightarrow{A} > 0$) we obtain

$$\lambda_1(A_{\uparrow}) < 0, \qquad \lambda_2(A_{\uparrow}) > 0. \tag{105}$$

This result will be discussed later. For the moment let us analyze the second boundary value problem. If the boundary condition at $x = 1$ is given by

$$y(t, 1) = y_1(t) \quad \text{for } 0 \leqslant t \leqslant T \tag{106}$$

we obtain new coefficients for partial differential Eq. (95):

$$S = S_{\downarrow} = \begin{bmatrix} -1/s & \\ & 1/s \end{bmatrix}, \qquad A = A_{\downarrow} = \frac{1}{s}\begin{bmatrix} -a_{22} & -a_{21} \\ a_{12} & a_{11} \end{bmatrix}. \tag{107}$$

Matrix $A_{\downarrow}$ has eigenvalues

$$\begin{aligned} \lambda_1(A_{\downarrow}) &= \frac{a_{11} - a_{22} - \sqrt{(a_{11} - a_{22})^2 + 4\det \overrightarrow{A}}}{2s} = -\lambda_2(A_{\uparrow}), \\ \lambda_2(A_{\downarrow}) &= \frac{a_{11} - a_{22} + \sqrt{(a_{11} - a_{22})^2 + 4\det \overrightarrow{A}}}{2s} = -\lambda_1(A_{\uparrow}) \end{aligned} \tag{108}$$

and due to the assumption that system (95), (97) is interior stable we obtain

$$\lambda_1(A_{\downarrow}) < 0, \qquad \lambda_2(A_{\downarrow}) > 0. \tag{109}$$

It is easy to verify that

$$\det A_{\uparrow} = \det A_{\downarrow} = -\frac{1}{s^2} \det \overrightarrow{A} < 0. \tag{110}$$

Moreover, the matrices $A_{\uparrow}$ and $A_{\downarrow}$ have one positive element on their diagonals: $-a_{11}/s > 0$ in the case of (103), and $-a_{22}/s > 0$ in the case of (107). This is contradictory to the assumptions of Theorem 1. Thus both analyzed boundary value problems: (95), (102) and (95), (106), are interior unstable. The eigenvalues of the matrices $A_{\uparrow}$ and $A_{\downarrow}$ are real and have different signs (see (104) and (108)). The negative eigenvalues suggest that the instability is not so strong as in the final-boundary value problem (95), (98). Indeed, computer simulations have given correct results for boundary value problems (95), (102) and (95), (106). These computer experiments suggest that stability is a sufficient but not a necessary condition for accurate computer calculations.

An advantage of calculation in both space directions is that it allows one to find a proper density of discretization experimentally. First, consider the boundary condition

$$y(t,0) = y_0(t) \quad \text{for } 0 \leqslant t \leqslant T, \tag{111}$$

where the final time T is sufficiently large ($T > 4s$). Now take the discretization Δt and apply the method of characteristics to obtain the numerical approximation of

$$y(t,1) \quad \text{for } 1/s \leqslant t \leqslant T - 1/s. \tag{112}$$

In the next step take (112) as a boundary condition for backward space calculations. Once more use the method of characteristics to obtain the approximation

$$y(t,0) \quad \text{for } 2/s \leqslant t \leqslant T - 2/s. \tag{113}$$

Comparing (in the sense of some metrics) the assumed values (111) and results (113) of the forward and backward computations, one obtains an error in the numerical calculations. These calculations can be repeated for a different Δt to obtain information about how the error and the time of computation depend on Δt.

4. Example

The method of an experimental choice of grid density was verified for the mathematical model of gas flow. Let us consider a system (see Fig. 3) which consists of a natural gas field, a pumping station, a long gas pipeline (e.g., two hundred kilometers in length and 0.7 meters in diameter) and gas consumer. Assume that the flow Q and the gas pressure P do not differ too much from their average values, so that the gas pipeline can be described by the linear hyperbolic equation

$$\frac{\partial}{\partial t}\begin{bmatrix} Q \\ P \end{bmatrix} + \begin{bmatrix} 0 & 1/M \\ M & 0 \end{bmatrix} \frac{\partial}{\partial x}\begin{bmatrix} Q \\ P \end{bmatrix} = \begin{bmatrix} -\rho & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} Q \\ P \end{bmatrix}, \tag{114}$$

where $\rho = \eta LM/D$ is the resistance of the gas flow and depends on Mach number M ($M \approx 0.01$), the length of pipeline L, the pipeline diameter D and the constant coefficient η. All the variables x, t, P, Q are normalized. The points $x = 0$ and $x = 1$ correspond respectively to the compressor end and the consumer end of the pipeline. The variable t is scaled with respect to the time in which the sonic wave

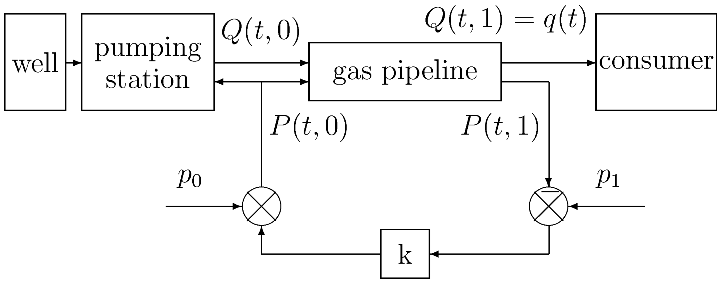

Fig. 3. The system of natural gas distribution with a pipeline described by a partial differential equation.

travels from the compressor end of the pipeline to the consumer end of the pipeline. The pressure P and flow Q are scaled with respect to the average pressure and flow.

The boundary conditions are given by equations

$$P(t,0) - p_0 = k\big(p_1 - P(t,1)\big), \tag{115}$$

$$Q(t,1) = q(t), \tag{116}$$

where p_0 is assumed to be a nominal gas pressure obtained by the compressor and p_1 is a nominal pressure of gas supplying the consumer. Eq. (115) describes a strategy of the control system. The gas pressure $P(t,1)$ at the consumer end of the pipeline is measured and compared with the nominal value p_1. The difference $p_1 - P(t,1)$ is multiplied by the feedback coefficient k and sent to the pumping station. The compressor control system keeps the gas pressure $P(t,0)$ at the level determined by (115) to compensate for the deviation of gas pressure at the consumer end of pipeline. The time-varying gas consumption $q(t)$ is included to the mathematical model by boundary condition (116). Due to the transformation

$$\begin{bmatrix} Q \\ P \end{bmatrix} = \begin{bmatrix} 1/\sqrt{M} & 1/\sqrt{M} \\ -\sqrt{M} & \sqrt{M} \end{bmatrix} y, \tag{117}$$

Eq. (114) takes the canonical form

$$\frac{\partial y}{\partial t} + S\frac{\partial y}{\partial x} = Ay, \tag{118}$$

where

$$S = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad A = -\frac{\rho}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Boundary conditions (115), (116) take the form

$$\begin{bmatrix} y^-(t,1) \\ y^+(t,0) \end{bmatrix} = B\begin{bmatrix} y^-(t,0) \\ y^+(t,1) \end{bmatrix} + \begin{bmatrix} u^e(t) \\ u^b(t) \end{bmatrix}, \tag{119}$$

where

$$B = \begin{bmatrix} 0 & -1 \\ 1 & -2k \end{bmatrix}, \qquad u^e(t) = \sqrt{M} q(t), \qquad u^b(t) = k\sqrt{M} q(t) + (p_0 + kp_1)/\sqrt{M}.$$

The progressive sonic wave $s^+ = 1$ and the return sonic wave $s^- = -1$ are normalized. The system considered is interior stable, but not asymptotically. Although elements on the diagonal of matrix A are negative, $\det A = 0$. This means that the energy of the gas is dissipated in the pipeline as usual, but energy is constant in the case where there is no gas flow, $Q(t, x) = q(t) = 0$, and the gas pressure is constant $P(t, x) = (p_0 + kp_1)/(1 + k)$ for all $0 \leqslant x \leqslant 1$ and $t \geqslant 0$. The energy functional which allows one to prove the interior stability has the form

$$E(t) = \int_0^1 y^{\mathrm{T}} \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} y \, \mathrm{d}t. \tag{120}$$

If $k = 0$ the considered system is boundary stable but not asymptotically. In the opposite case ($k \neq 0$), system (118), (119) is boundary unstable. The feedback introduced by the control system leads to the destabilization of the boundary condition (119).

The numerical method presented in Section 2 was used to find the solution of Eq. (118) with the non-homogeneous boundary condition (119). The discrete equation (93) takes the form

$$y_{j+1} = \Phi y_j + U_0 u^b_{j+1} + U_1 u^e_{j+1}, \tag{121}$$

where

$$U_0 = [0\ 1\ 0\ \dots\ 0]^{\mathrm{T}} \quad \text{and} \quad U_1 = [0\ \dots\ 0\ 1\ 0]^{\mathrm{T}}.$$

Computer calculations were performed for the following parameters: $\rho = 20$, $M = 0.01$, $k = 0.5$, $p_0 = 1.1$, $p_1 = 0.9$ and the final time $T = 25$. The initial conditions

$$\begin{matrix} Q(0, x) = 1 \\ P(0, x) = 1.1 - \rho M x \end{matrix} \quad \text{for } 0 \leqslant x \leqslant 1 \tag{122}$$

were imposed. The gas consumption was approximated by the function

$$q(t) = 1 + 0.1 \sin(0.2t) - 0.01 \sin(t) \quad \text{for } 0 \leqslant t \leqslant T. \tag{123}$$

With these assumptions the initial-boundary value problem (114)–(116) was solved numerically and the numerical solution at the boundary $x = 0$

$$\begin{matrix} Q(t, 0) \\ P(t, 0) \end{matrix} \quad \text{for } 0 \leqslant t \leqslant T \tag{124}$$

was stored in the computer memory. Then these data were used to find numerically the solution at the boundary $x = 1$:

$$\begin{matrix} Q(t, 1) \\ P(t, 1) \end{matrix} \quad \text{for } 1 \leqslant t \leqslant T - 1. \tag{125}$$

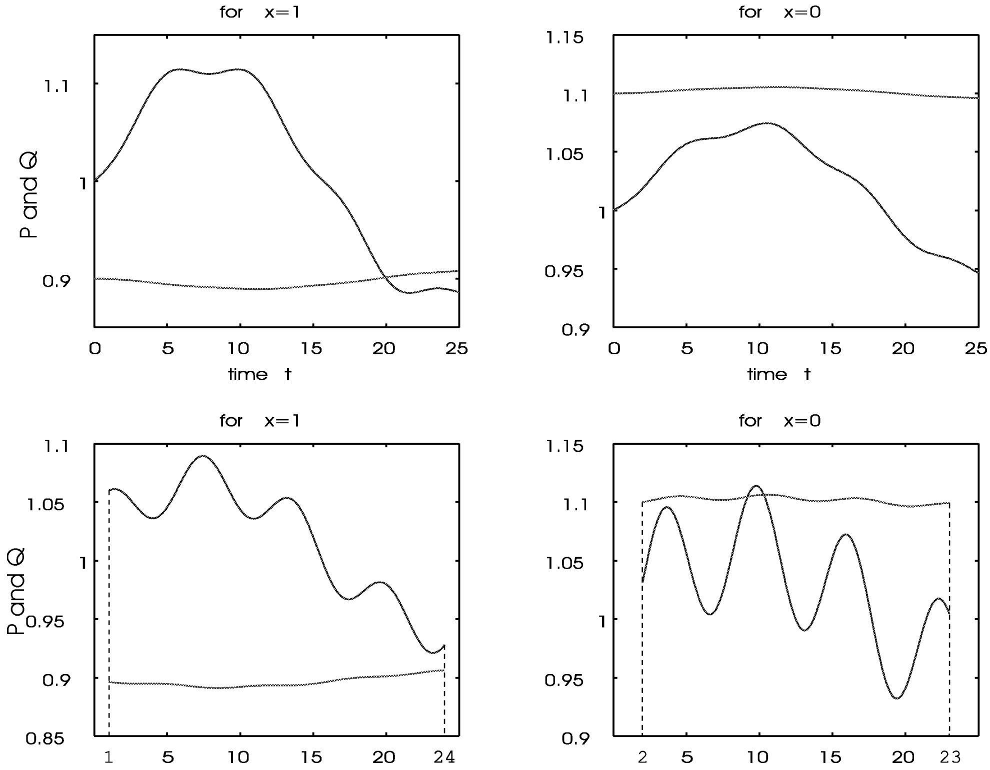

Fig. 4. The numerical solution of the initial-boundary value problem (upper plots) and the boundary value problem (lower plots). Considerable differences between solutions result from too large a density of discretization ($\Delta x = \frac{1}{10}$).

Finally, the calculations were repeated in the opposite space direction and the discrete approximation of

$$\begin{matrix} Q^*(t,0) \\ P^*(t,0) \end{matrix} \quad \text{for } 1 \leqslant t \leqslant T-2 \tag{126}$$

was obtained. Comparing the values in (124) with result (126), obtained after first forward and then backwards in space calculations, the error of the numerical computation was calculated for the variables Q and P individually:

$$\begin{aligned} \Delta Q &= \left[\frac{1}{T-4}\int_2^{T-2}\bigl(Q(t,0)-Q^*(t,0)\bigr)^2\,\mathrm{d}t\right]^{0.5}, \\ \Delta P &= \left[\frac{1}{T-4}\int_2^{T-2}\bigl(P(t,0)-P^*(t,0)\bigr)^2\,\mathrm{d}t\right]^{0.5}. \end{aligned} \tag{127}$$

To observe the degree of error dependence on the density of discretization, similar calculations were repeated for the nine cases: $\Delta x = \frac{1}{5}, \frac{1}{10}, \ldots, \frac{1}{45}$. The results are presented in Figs. 4–6. On the upper part

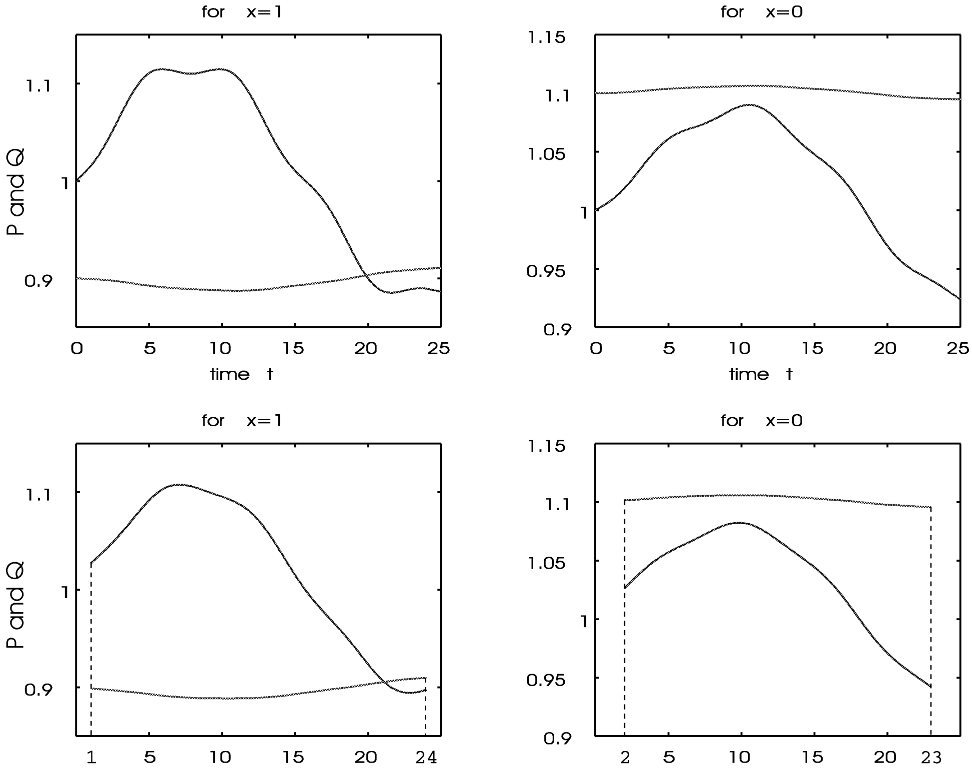

Fig. 5. The numerical solution of the initial-boundary value problem (upper plots) and the boundary value problem (lower plots). Small differences between solutions result from an optimal density of discretization ($\Delta x = \frac{1}{30}$).

of Fig. 4 the numerical solution of the initial-boundary value problem (114)–(116), (122) is presented. The boundary condition (124) for the forward in space calculations is presented in the right upper part of Fig. 4. Solution (125) obtained for $x = 1$ is presented in the lower left part. Result (126) of backwards in space calculations is presented in the lower right part of Fig. 4. These results were obtained for the space discretization $\Delta x = \frac{1}{10}$. The comparison of plots from the right hand side of Fig. 4 leads to the conclusion that too large a discretization makes the error considerable. Fig. 5 shows results for the space discretization $\Delta x = \frac{1}{30}$ which is three times smaller. In this case the accuracy is satisfactory. Fig. 6 shows the dependence of error (127) and the time of the calculations on the grid density. All these plots suggest that the time of calculations increases rapidly when discretization is refined. On the other hand, the error dependence on the grid discretization is insignificant for the small discretization. These observations lead to the conclusion that $\Delta x = \frac{1}{30}$ is the optimal discretization in the case considered.

5. Conclusions

The main assumptions of Theorems 1 and 2 are the Hurwitz stability of matrix A and the Schur stability of matrix B. It is right to suppose that the other assumptions result from restrictions which

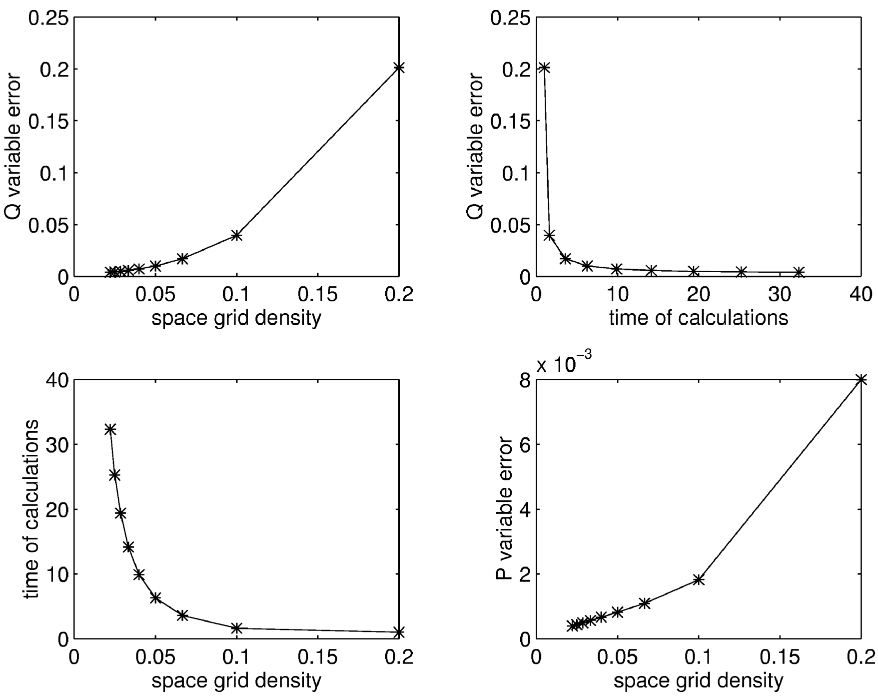

Fig. 6. Dependencies between accuracy, time of computations and density of discretization Δx.

led to the diagonal form of matrix G. There are two possibilities of avoiding this disadvantage. The first one is to put additional restrictions on matrices A and B. The D-symmetribility of these matrices (see [8], [9]) makes energy functional (11) much more useful in verifying the stability of hyperbolic systems. The alternative is to seek other Lyapunov functionals. Unfortunately, how to do it is still an open question. Proposition 2 demonstrates a well understood rule that the reversal of time direction transforms a stable system into an unstable one. The instability of the continuous final-boundary value problem leads to the numerical instability of the corresponding discrete system. Computational experiments have showed that this is a strong instability. The large errors observed led to a computing overflow and make it impossible to integrate backwards the partial differential equation (95). The computer modeling of real physical systems points out that it is possible to use the method of characteristics in both directions of the space variable but only in one direction of the time variable. An advantage of calculation in both space directions is that it allows one to find a proper density of discretization experimentally.

Appendix. Justification for Conjecture 1

Following Proposition 1 it is enough to prove that

$$\left\| G_1^{1/2} \Phi_b G_2^{-1/2} \right\| < 1, \tag{128}$$

where $\dim G_1 = 2(1+1/\Delta x) \times 2(1+1/\Delta x)$ and $\dim G_2 = 2/\Delta x \times 2/\Delta x$. Inequality (128) is satisfied if all the eigenvalues of the matrix

$$\Theta_b = G_2^{-1/2} \Phi_b^{\mathrm{T}} G_1 \Phi_b G_2^{-1/2} \tag{129}$$

are within the segment $[0, 1)$. To verify this let us consider the determinant

$$\Xi_J = \det(\Theta_b - \lambda I) = \det \begin{bmatrix} d & e & 0 & 0 & \dots & 0 & 0 & f & h \\ e & k & 0 & 0 & \dots & 0 & 0 & l & m \\ 0 & 0 & p & r & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & r & u & \dots & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & p & r & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & r & u & 0 & 0 \\ f & l & 0 & 0 & \dots & 0 & 0 & v & w \\ h & m & 0 & 0 & \dots & 0 & 0 & w & z \end{bmatrix}, \tag{130}$$

where

$$\begin{aligned}
&d = (1+a_{11}\Delta t)^2\big(1+b_{11}^2+gb_{21}^2\big)+ga_{21}^2\Delta t^2-\lambda,\\
&e = a_{21}\Delta t(1+a_{22}\Delta t)\sqrt{g}+a_{12}\Delta t(1+a_{11}\Delta t)\big(b_{21}^2\sqrt{g}+1/\sqrt{g}+b_{11}^2/\sqrt{g}\big),\\
&f = a_{21}\Delta t(1+\Delta t a_{11})(b_{11}b_{12}+gb_{21}b_{22}),\\
&h = (1+\Delta t a_{11})(1+\Delta t a_{22})\big(b_{11}b_{12}/\sqrt{g}+b_{21}b_{22}\sqrt{g}\big),\\
&k = (1+a_{22}\Delta t)^2+a_{12}^2\Delta t^2\big(b_{21}^2+1/g+b_{11}^2/g\big)-\lambda,\\
&l = a_{12}a_{21}\Delta t^2\big(b_{21}b_{22}\sqrt{g}+b_{11}b_{12}/\sqrt{g}\big),\\
&m = a_{12}\Delta t(1+a_{22}\Delta t)(b_{21}b_{22}+b_{11}b_{12}/g),\\
&p = a_{21}^2\Delta t^2 g+(1+a_{11}\Delta t)^2-\lambda,\\
&r = a_{12}\Delta t(1+a_{11}\Delta t)/\sqrt{g}+a_{21}\Delta t(1+a_{22}\Delta t)\sqrt{g},\\
&u = a_{12}^2\Delta t^2/g+(1+a_{22}\Delta t)^2-\lambda,\\
&v = (1+a_{11}\Delta t)^2+a_{21}^2\Delta t^2\big(b_{12}^2+g+b_{22}^2 g\big)-\lambda,\\
&w = a_{12}\Delta t(1+a_{11}\Delta t)/\sqrt{g}+a_{21}\Delta t(1+a_{22}\Delta t)\big(b_{12}^2/\sqrt{g}+\sqrt{g}+b_{22}^2\sqrt{g}\big),\\
&z = (1+a_{22}\Delta t)^2\big(1+b_{22}^2+b_{12}^2/g\big)+a_{12}^2\Delta t^2/g-\lambda, \qquad \dim \Xi_J = 2/\Delta x \times 2/\Delta x.
\end{aligned} \tag{131}$$

Determinant (130) is a product of the two polynomials

$$\Xi_J = \Xi_i \Xi_b, \tag{132}$$

where

$$\Xi_i = \left(pu - r^2\right)^{-4+2/\Delta x}, \tag{133}$$

$$\Xi_b = \det \begin{bmatrix} d & e & f & h \\ e & k & l & m \\ f & l & v & w \\ h & m & w & z \end{bmatrix}. \tag{134}$$

The Schur stability of (133) follows, in a similar way to (68), from conditions (57) since $pu - r^2 = ac - b^2$. Unfortunately, it is difficult to prove that the Schur stability of (134) is a consequence of the assumptions of Theorems 1 and 2 for sufficiently small Δt. Anyway, the presented conjecture is useful in practice and enables us to verify the numerical stability for each individual case. There are no difficulties with numerical allocation of zeros of polynomial (134).

References

[1] R. Courant, D. Hilbert, Methods of Mathematical Physics, Interscience, New York, 1962.

[2] J. Edenhofer, G. Schmitz, An implicit method of characteristics for the solution of hyperbolic initial-boundary-value problems and its convergence, Computing 26 (1981) 257–264.

[3] M.D. Gunzburger, On the stability of Galerkin methods for initial-boundary value problem for hyperbolic systems, Math. Comp. 31 (1977) 661–675.

[4] L. Ioffe, S.S. Pinter, Applying asynchronous parallel characteristic methods for solving systems of hyperbolic PDEs, J. Sci. Comput. 8 (1993) 195–218.

[5] R.A. Nicolaides, Analysis and convergence of the MAC scheme. I. The linear problem, SIAM J. Numer. Anal. 29 (1992) 1579–1591.

[6] K.H. Parker, C.J.H. Jones, Forward and backward running waves in the arteries: Analysis using the method of characteristics, Trans. ASME J. Biomech. Engrg. 112 (1990) 322–326.

[7] D.L. Simpson, A Numerical Method of Characteristics for Solving Hyperbolic Partial Differential Equations, Michigan University Press, Ann Arbor, 1967.

[8] M. Ziółko, Application of Lyapunow functionals to studying stability of linear hyperbolic systems, IEEE Trans. Automat. Control 10 (1990) 1173–1176.

[9] M. Ziółko, S. Białas, Robust stability of symmetrizable hyperbolic systems, submitted to Linear Algebra.