Algorithms for the exact solution of two matrix equations occurring in the driven cavity problem

OCEANOLOGIA, 41 (1), 1999.
pp. 3–23.

KEYWORDS
Navier-Stokes equations
Elimination of pressure
Driven cavity problem
Improvement of the
process of solution

Maria Ewa Klonowska,
Włodzimierz Juliusz Prosnak
Institute of Oceanology,
Polish Academy of Sciences,
Powstańców Warszawy 55, PL–81–712 Sopot, Poland;
e-mail: klonowsk@iopan.gda.pl

Manuscript received December 10, 1998, in final form January 20, 1999.

Abstract

It is shown that two particular systems of linear equations, derived in an earlier paper by Prosnak & Kosma (1991), can be solved in an exact time- and storage-saving manner. First of all, by the proper elimination of unknowns, each system can be reduced to a smaller one containing only half of the unknowns. In the first case, the matrix of coefficients of the so reduced system turns out to be tridiagonal, its elements consisting of square submatrices. Moreover, the reduced system can be split into two independent ones. In the second case, the matrix of the reduced system can be presented as the product of two triangular ones, each one being partitioned in square submatrices. Corresponding algorithms and computer programs have been developed in order to investigate whether some economy in storage and computing time is really attainable. Affirmative conclusions are drawn from the results of computations. This means that the new method of solving problems governed by the Navier-Stokes equations, presented in the cited paper, can be applied in a more effective manner.

1. Introduction

An important direction of research in contemporary numerical fluid dynamics concerns methods for the determination of unsteady, viscous and incompressible flows governed by the system of equations referred to for the sake of convenience as the Navier-Stokes equations, although they contain also the continuity equation.

For obvious reasons, the properties of these newly developed methods are usually investigated in the *plane* case, when the Navier-Stokes equations reduce to the following form:

$$\left.\begin{aligned} &u_x + v_y = 0;\\ &u_t + uu_x + vu_y = -p_x/\rho + \nu\Delta u;\\ &v_t + uv_x + vv_y = -p_y/\rho + \nu\Delta v \end{aligned}\right\} \tag{1}$$

in a rectilinear system of co-ordinates x, y. The symbols t, Δ, ρ, ν denote time, Laplace operator, constant density and constant kinematic viscosity of the fluid, respectively, the indices referring to partial derivatives.

For the sake of completeness and clarity it should be recalled that the unknown functions in (1), *i.e.* the velocity components

$$u = u\,(x, y, t), \qquad v = v\,(x, y, t) \tag{2}$$

in the direction of the x and y axes respectively, and the pressure

$$p = p\,(x, y, t) \tag{3}$$

are 'incompatible' as far as boundary conditions are concerned. Namely, such conditions have to be imposed on all boundaries of the domain of solution for the components (2), but at just one point of this domain for the pressure (3). Moreover, no derivative of pressure with respect to time appears in (1).

It should be recalled that the 'incompatibility' can be retraced to the oversimplified physical model of flow described by the system (1), in particular – to the assumption of constant density.

In order to avoid the inconveniences and problems arising out of this 'incompatibility', and to reduce the number of unknown functions, the majority of methods of solution start with the elimination of pressure from the system (1). Differentiation is usually applied for this purpose which, however, increases the order of the system. Consequently, boundary conditions for *derivatives* of components (2) have to be introduced, although such conditions do not appear in the original problem. Nevertheless, these artificial, 'unphysical' conditions have to be defined somehow. They are always disputable, and can be validated only *a posteriori*, and indirectly, by a comparison of the results of applying different methods to the same problem. For more than 30 years now, the so-called *driven cavity problem* has usually been used as such a *test problem.* This is also the case in the present considerations.

A new method for determining plane, unsteady flows of viscous, incompressible fluids was presented by Prosnak & Kosma (1991), the novelty consisting in the elimination of pressure from the governing system of equations by means of the *condition of univalence*. This condition was expressed in the form of the following requirement:

$$\oint_l (p_x\,dx + p_y\,dy) = 0; \qquad (t = \text{const}), \tag{4}$$

imposed at any fixed time t on the total differential of pressure. The derivatives in (4) stem from the Navier-Stokes equations (1), and the symbol l denotes any contour in the domain of solution, contractible to a regular point of the domain.

Because *integration*, and not the customary *differentiation*, is applied to eliminate pressure, no need arises for artificial boundary conditions to be imposed on *derivatives* of the velocity components (2). This represents the fundamental virtue of the new method.

This method was applied by Prosnak & Kosma (1991) and by Klonowska & Kołodziejczyk (in press) to the *driven cavity problem* (Fig. 1). The main features of this problem must, however, be presented here in order to avoid misunderstandings.

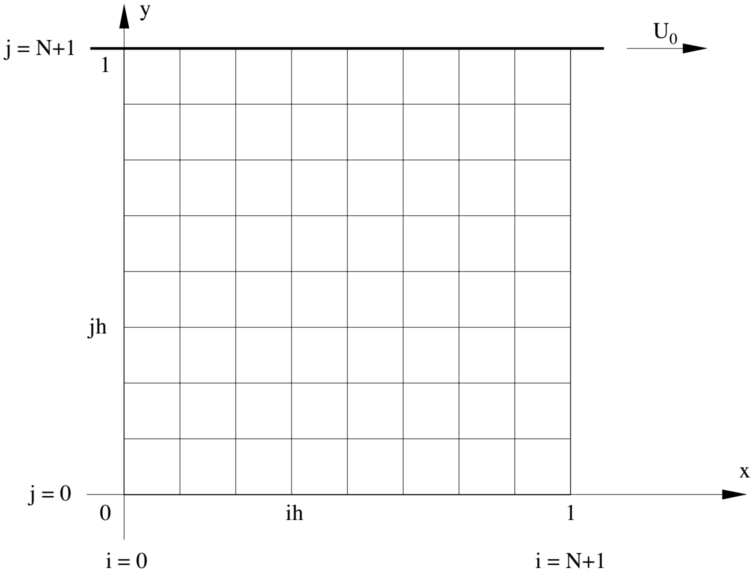

Fig. 1. The ‘driven cavity’ problem (Prosnak & Kosma 1991)

The square cavity containing the viscous fluid is bounded by three fixed walls

$$y = 0; \qquad x \in [0,1], \tag{5}$$

$$x = 0; \qquad y \in [0,1], \tag{6}$$

$$x = 1; \qquad y \in [0, 1] \tag{7}$$

and by a fourth one

$$y = 1; \qquad x \in [0, 1], \tag{8}$$

which shifts in the x – direction with a given velocity

$$U_o = U_o\,(t), \tag{9}$$

as shown in Fig. 1.

The *boundary conditions* express the *impermeability* of all four walls, and the *non-slip* property of the fluid.

The *initial conditions* stem from the assumption that the fluid is at rest for $t \le 0$.

Function (9) has to comply with this condition.

We applied the finite difference method and the method of splines to the solution of the driven cavity problem in the two references cited. Hence, a quadratic grid was assumed in the domain of solution, the spacing of the grid being $h = 1/(N+1)$ depending on the integer N. The nodes i, j $\in$ [1, N] of the grid will be referred to as *internal* ones. Only at these nodes are the velocity components (2) unknown. At the other ones, distributed on the walls, they follow from the boundary conditions.

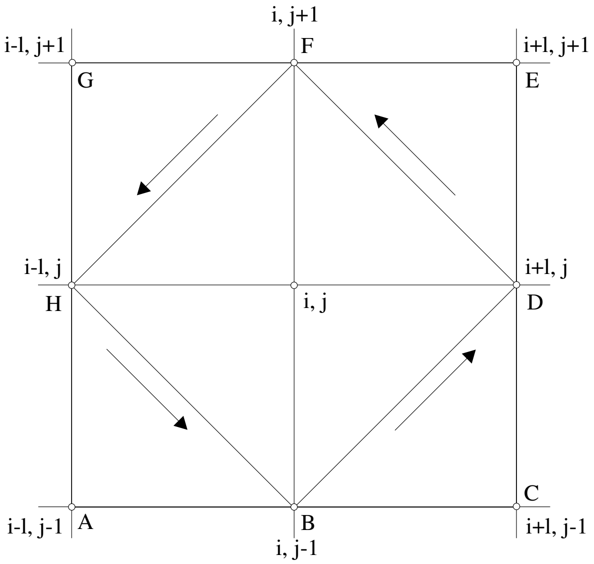

Fig. 2. A subdomain of the domain of solution containing two alternative contours of integration: BDFHB and ABCDEFGHA (Prosnak & Kosma 1991)

Now it should be well understood that the finite difference method and the method of splines concern solely the space variables x, y, the respective derivatives being replaced by differences of the functions of time

$$u_{ij} = u_{ij}(t), \qquad v_{ij} = v_{ij}(t); \qquad i, j \in [1,\, N] \tag{10}$$

corresponding to the internal nodes. The number of such functions is

$$2N^2. \tag{11}$$

The derivatives of these functions with respect to time, contained in the Navier-Stokes equations (1), remain. However, after discretisation of the domain of solution they refer solely to the nodes.

In such a manner the driven cavity problem was reduced within the framework of the method under consideration to a system of ordinary differential equations of the first-order containing derivatives of the velocity components (10) at the indicated discrete N^2 points of the domain of solution.

This system depends on three factors:

- on the alternative, whether the continuity equation is taken in its differential or its integral form;
- on the formulae approximating the spatial derivatives of the velocity components (2);
- on the contour of integration (see Fig. 2) adopted in the integrals (4) and – finally – in the integrals expressing the continuity condition.

In the first of the references cited, three cases were distinguished, which lead to the following matrix equations:

$$\mathbf{A}\dot{\mathbf{Y}} = \mathbf{W}, \tag{12}$$

$$\mathbf{B}\dot{\mathbf{Y}} = \mathbf{V}, \tag{13}$$

$$\mathbf{A}\dot{\mathbf{Y}} = \mathbf{U}. \tag{14}$$

In these equations the symbols $\mathbf{A}$ and $\mathbf{B}$ denote square constant matrices of the order $2N^2$, corresponding to the number (11) of unknown functions (10). The symbol $\dot{\mathbf{Y}}$ denotes the column vector

$$\dot{\mathbf{Y}} = [\dot{Y}_1,\, \dot{Y}_2, ...,\, \dot{Y}_m]^T; \qquad m = 2N^2, \tag{15}$$

consisting of the unknown derivatives of functions (10) with respect to time. The symbols $\mathbf{W}$, $\mathbf{V}$, $\mathbf{U}$ denote known column vectors, depending on the instantaneous *values* of functions (10) and on the kinematic viscosity ν.

The correspondence between the elements of the vectors $\mathbf{W}$, $\mathbf{V}$, $\mathbf{U}$ and the velocity components (10) will be not needed in the present considerations.

System (12) corresponds to the case when the continuity equation is taken in its *differential* form, the spatial derivatives of velocity components (2) are approximated by finite differences, and the contours of the type BDFHB (see Fig. 2) appear in the integrals (4).

System (13) corresponds to the case when the continuity equation is taken in its *integral* form, the spatial derivatives are approximated by the same finite differences, and the contours ABCDEFGHA (see Fig. 2) appear in the integrals (4) as well as in the integrals expressing the continuity equation.

Finally, system (14) corresponds to the case when the spatial derivatives of (2) are approximated by formulae stemming from the theory of cubic splines, the continuity equation is taken in its differential form, and contours of the type BDFHB (Fig. 2) appear in the integrals (2).

Systems (12) and (14) differ solely with respect to their right hand sides, so that further considerations will be confined to systems (12) and (13).

By introducing the respective inverses $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$ of the matrices of coefficients $\mathbf{A}$, $\mathbf{B}$, one can reduce (12) and (13) to the Cauchy normal form

$$\dot{\mathbf{Y}} = \mathbf{A}^{-1}\,\mathbf{W}, \tag{16}$$

$$\dot{\mathbf{Y}} = \mathbf{B}^{-1}\,\mathbf{V}, \tag{17}$$

where the inverses are constant.

In such a manner the solution of the driven cavity problem has been reduced to the *initial problem* for the set of $2N^2$ ordinary differential equations of the first-order, the unknown function of time $\mathbf{Y}$ representing the velocity components (10).

Any self-starting numerical method for the solution of any initial problem of the type (12), (13) or (14), exemplified by the Runge-Kutta method (see *e.g.* Lambert (1973)) makes use of the values of the vector of derivatives (15). Elements of this vector have to be evaluated at least at the initial and the final point of every time step. In the case of the Runge-Kutta method of an order larger than 1, such evaluations concern also a number of points situated within the time step. Hence, it is clear that the economy of such a method of integration applied to the problem under consideration depends directly on the economy of computing the vector (15) from eq. (12) or (13).

The most obvious method for evaluating this vector stems from formulae (16) and (17), and utilises the inverses of the matrices of coefficients. This is a very elegant method; however, it does require computation and storage of the inverses. Moreover, the number of multiplications of the order

$$2\,(N^2)^3 \tag{18}$$

and, consequently, the computer time involved in such an evaluation of the vector $\dot{\mathbf{Y}}$ as well as the corresponding storage may be quite large. These two aspects may very severely restrict application of this method.

Therefore, the vitally important question arises whether this vector could be evaluated more economically – by means of special algorithms making use of the particular properties of the matrices **A** and **B**.

The present considerations refer to the development and the comparative investigation of such algorithms.

2. The equation $\mathbf{A}\dot{\mathbf{Y}} = \mathbf{W}$

Introducing after Prosnak & Kosma (1991) an auxiliary tridiagonal submatrix ${\alpha_1}^{[N]}$ of the order N, and the identity submatrix $I^{[N]}$ of the same order

$$ {\alpha_1}^{[N]} = \begin{bmatrix} 0 & 1 & & & & 0 \\ -1 & 0 & 1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & -1 & 0 & 1 \\ 0 & & & & -1 & 0 \end{bmatrix}, \tag{19} $$

$$ I^{[N]} = \begin{bmatrix} 1 & & & & & 0 \\ & 1 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & 1 & \\ 0 & & & & & 1 \end{bmatrix}, \tag{20} $$

one can rewrite the matrix system (12) of ordinary differential equations in the following manner:

$$ \left. \begin{array}{l} \alpha_1 \dot{Y}_1 \qquad\qquad + \dot{Y}_{N+2} = W_1; \\ \alpha_1 \dot{Y}_2 - \dot{Y}_{N+1} + \dot{Y}_{N+3} = W_2; \\ \alpha_1 \dot{Y}_3 - \dot{Y}_{N+2} + \dot{Y}_{N+4} = W_3; \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_1 \dot{Y}_{N-1} - \dot{Y}_{2N-2} + \dot{Y}_{2N} = W_{N-1}; \\ \alpha_1 \dot{Y}_N - \dot{Y}_{2N-1} \qquad\qquad = W_N, \end{array} \right\} \tag{21} $$

$$\left.\begin{aligned} -\dot{Y}_2 + \alpha_1\dot{Y}_{N+1} &= W_{N+1};\\ \dot{Y}_1 - \dot{Y}_3 + \alpha_1\dot{Y}_{N+2} &= W_{N+2};\\ \dot{Y}_2 - \dot{Y}_4 + \alpha_1\dot{Y}_{N+3} &= W_{N+3};\\ \dots\dots\dots\dots\dots&\dots\dots\dots\dots\\ \dot{Y}_{N-2} - \dot{Y}_N + \alpha_1\dot{Y}_{2N-1} &= W_{2N+1};\\ \dot{Y}_{N-1} \qquad\quad + \alpha_1\dot{Y}_{2N} &= W_{2N}, \end{aligned}\right\} \tag{22}$$

the superscript at ${\alpha_1}^{[N]}$ and at the identity matrix $I^{[N]}$ being omitted for the sake of simplicity.

It should be recalled (Prosnak & Kosma 1991) that the determinant of the submatrix (19) vanishes for odd values of N. This peculiar property follows from the difference formulae applied for approximation of the spatial derivatives of (2), but it has no other consequences besides confining the computations to even numbers of nodes

$$N = 2\,P, \tag{23}$$

where P denotes an odd or an even integer.

All the terms in (21)–(22) are column vectors of the order N; they represent elements of column vectors of the order $2N^2$ appearing in (12):

$$\dot{\mathbf{Y}} = [\dot{Y}_1, \dot{Y}_2, ..., \dot{Y}_{2N}]^T, \tag{24}$$

$$\mathbf{W} = [W_1, W_2, ..., W_{2N}]^T. \tag{25}$$

On the other hand, they depend on the real elements of these vectors, such as those introduced in (15) with respect to $\dot{\mathbf{Y}}$. In the case of the vectors of unknowns $\dot{Y}_k$ appearing in (24), the appropriate relations can be written as follows:

$$\left.\begin{aligned} &\dot{Y}_k = \left[\dot{y}_{1+(k-1)N}, \dot{y}_{2+(k-1)N}, ..., \dot{y}_{i+(k-1)N}, ..., \dot{y}_{N+(k-1)N}\right]^T;\\ &k \in [1,\ 2N], \qquad\qquad\qquad i \in [1,\ N]. \end{aligned}\right\} \tag{26}$$

The analogous relations concern the vectors

$$W_k; \qquad k \in [1,\ 2N]. \tag{27}$$

The structure of the matrix of coefficients in (21)–(22) is fairly regular, as can be seen in Fig. 3, which exemplifies this matrix for the case $N = 8$.

The symbols α_1 and I in this Figure denote the N-th order submatrices (19) and (20), the superscripts at α_1 and I being, however, omitted again. The superscript is retained only by the symbol of matrix $\mathbf{A}$ – see Fig. 3.

$$\mathbf{A}^{[8]} = \begin{bmatrix}
\alpha_1 & & & & & & & & & \mathrm{I} & & & & & & \\
& \alpha_1 & & & & & & & \mathrm{-I} & & \mathrm{I} & & & & & \\
& & \alpha_1 & & & & & & & \mathrm{-I} & & \mathrm{I} & & & & \\
& & & \alpha_1 & & & & & & & \mathrm{-I} & & \mathrm{I} & & & \\
& & & & \alpha_1 & & & & & & & \mathrm{-I} & & \mathrm{I} & & \\
& & & & & \alpha_1 & & & & & & & \mathrm{-I} & & \mathrm{I} & \\
& & & & & & \alpha_1 & & & & & & & \mathrm{-I} & & \mathrm{I} \\
& & & & & & & \alpha_1 & & & & & & & \mathrm{-I} & \\
& \mathrm{-I} & & & & & & & \alpha_1 & & & & & & & \\
\mathrm{I} & & \mathrm{-I} & & & & & & & \alpha_1 & & & & & & \\
& \mathrm{I} & & \mathrm{-I} & & & & & & & \alpha_1 & & & & & \\
& & \mathrm{I} & & \mathrm{-I} & & & & & & & \alpha_1 & & & & \\
& & & \mathrm{I} & & \mathrm{-I} & & & & & & & \alpha_1 & & & \\
& & & & \mathrm{I} & & \mathrm{-I} & & & & & & & \alpha_1 & & \\
& & & & & \mathrm{I} & & \mathrm{-I} & & & & & & & \alpha_1 & \\
& & & & & & \mathrm{I} & & & & & & & & & \alpha_1
\end{bmatrix}$$

Fig. 3. An example – for $N = 8$ – of the matrix of coefficients in (12), partitioned into submatrices of the order N

It should be understood that the empty squares in Fig. 3 are in fact occupied by null submatrices of the order $N = 8$.

The matrix $\mathbf{A}^{[N]}$ is not tridiagonal. However, by virtue of its regular structure, half of the unknowns, *i.e.*

$$\dot{Y}_k; \qquad k \in [1,\, N] \tag{28}$$

can be easily eliminated, which leads to a 'reduced' system of equations, containing just the remaining unknowns

$$\dot{Y}_k; \qquad k \in [N+1,\, 2N]. \tag{29}$$

For this purpose the upper half of eqs. (21)–(22) is multiplied by the inverse α_1^{-1}, which yields the following formulae:

$$\left.\begin{array}{l}
\dot{Y}_1 = -\alpha_1^{-1}\dot{Y}_{N+2} \qquad\qquad\qquad + \alpha_1^{-1}W_1; \\
\dot{Y}_2 = -\alpha_1^{-1}\dot{Y}_{N+3} + \alpha_1^{-1}\dot{Y}_{N+1} + \alpha_1^{-1}W_2; \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\dot{Y}_{N-1} = -\alpha_1^{-1}\dot{Y}_{2N} + \alpha_1^{-1}\dot{Y}_{2N-2} + \alpha_1^{-1}W_{N-1}; \\
\dot{Y}_N = \qquad\qquad\qquad \alpha_1^{-1}\dot{Y}_{2N-1} + \alpha_1^{-1}W_N.
\end{array}\right\} \tag{30}$$

Substitution of these formulae into equations representing the lower half of (21)–(22) results – after some rather obvious transformations – in the following 'reduced system':

$$\left.\begin{aligned}
b\dot{Y}_{N+1} + \dot{Y}_{N+3} &= R_{N+1};\\
a\dot{Y}_{N+2} + \dot{Y}_{N+4} &= R_{N+2};\\
\dot{Y}_{N+1} + a\dot{Y}_{N+3} + \dot{Y}_{N+5} &= R_{N+3};\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\dot{Y}_{2N-4} + a\dot{Y}_{2N-2} + \dot{Y}_{2N} &= R_{2N-2};\\
\dot{Y}_{2N-3} + a\dot{Y}_{2N-1} &= R_{2N-1};\\
\dot{Y}_{2N-2} + b\dot{Y}_{2N} &= R_{2N},
\end{aligned}\right\} \tag{31}$$

which has a *tridiagonal matrix of coefficients*. The newly introduced symbols denote

$$a = \alpha_1^2 - 2I; \qquad b = \alpha_1^2 - I, \tag{32}$$

$$\left.\begin{aligned}
R_{N+1} &= \alpha_1 W_{N+1} + W_2;\\
R_{N+2} &= \alpha_1 W_{N+2} + W_3 - W_1;\\
\ldots\ldots&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
R_{2N-1} &= \alpha_1 W_{2N-1} + W_N - W_{N-2};\\
R_{2N} &= \alpha_1 W_{2N} - W_{N-1}.
\end{aligned}\right\} \tag{33}$$

It is evident that the system (31) consists in fact of two independent ones. The first one contains unknowns with *odd* subscripts and can be rewritten as

$$\left.\begin{aligned}
b\dot{Y}_{N+1} + \dot{Y}_{N+3} &= R_{N+1};\\
\dot{Y}_{N+1} + a\dot{Y}_{N+3} + \dot{Y}_{N+5} &= R_{N+3};\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\dot{Y}_{2N-3} + a\dot{Y}_{2N-1} &= R_{2N-1},
\end{aligned}\right\} \tag{34}$$

while the second one contains unknowns provided with *even* subscripts, and has the following form:

$$\left.\begin{aligned}
a\dot{Y}_{N+2} + \dot{Y}_{N+4} &= R_{N+2};\\
\dot{Y}_{N+2} + a\dot{Y}_{N+4} + \dot{Y}_{N+6} &= R_{N+4};\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\dot{Y}_{2N-4} + a\dot{Y}_{2N-2} + \dot{Y}_{2N} &= R_{2N-2};\\
\dot{Y}_{2N-2} + b\dot{Y}_{2N} &= R_{2N}.
\end{aligned}\right\} \tag{35}$$

Both can be solved by a quite well-known process consisting of two operations: the elimination of successive unknowns (the ‘downward’ part of the process), and the evaluation of consecutive unknowns, starting with the one corresponding to the maximum value of the subscript (the ‘upward’ part of the process).

Simple recursive formulae are characteristic of this approach.

In the case of system (34) they are derived as follows. The first equation in (34) is solved with respect to $\dot{Y}_{N+1}$, yielding

$$\begin{aligned}\dot{Y}_{N+1} &= -b^{-1}\dot{Y}_{N+3} + b^{-1}R_{N+1} = \\ &= -q_1\dot{Y}_{N+3} + u_1. \end{aligned}\tag{36}$$

This result is substituted in the second equation in (34), which is then solved with respect to $\dot{Y}_{N+3}$:

$$\begin{aligned}\dot{Y}_{N+3} &= -(-q_1 + a)^{-1}\dot{Y}_{N+5} + (-q_1 + a)^{-1}(R_{N+3} - u_1) = \\ &= -q_2\dot{Y}_{N+5} + u_2. \end{aligned}\tag{37}$$

This operation can be repeated with respect to consecutive equations in (34), right down to the last one. Hence, the *first part* of the process consists in the recursive evaluation of the submatrices

$$\left.\begin{aligned} &q_i = (-q_{i-1} + a)^{-1}; \qquad i \in [1,\, N/2], \\ &q_1 = b^{-1} \end{aligned}\right\}\tag{38}$$

and of the column vectors

$$\left.\begin{aligned} &u_i = q_i(R_{N+2i-1} - u_{i-1}); \qquad i \in [1,\, N/2], \\ &u_o = 0. \end{aligned}\right\}\tag{39}$$

They appear in the general relation

$$\dot{Y}_{N+2i-1} = -q_i\dot{Y}_{N+2i+1} + u_i; \qquad i \in [1,\, N/2]\tag{40}$$

connecting two consecutive vectorial unknowns. The last relation, *i.e.* the one corresponding to $i = N/2$, reduces to the simple equation

$$\dot{Y}_{2N-1} = u_{N/2}\tag{41}$$

because the unknown $\dot{Y}_{2N+1}$ does not exist. Consequently, *the second part* of the process under consideration consists in evaluating the unknowns by the use of the relations (40), starting with the unknown (41).

It should be noted that the inverses (38) do not depend on the right hand sides in (35), so that they can be evaluated and stored with a view to saving computing time.

In the case of system (35) the same approach is applied; however, slightly different recursive formulae appear:

$$\begin{aligned}\dot{Y}_{N+2} &= -a^{-1}\dot{Y}_{N+4} + a^{-1}R_{N+2} = \\ &= -q_1\dot{Y}_{N+4} + u_1; \end{aligned}$$

$$\begin{aligned} \dot{Y}_{N+4} &= -(-q_1+a)^{-1}\dot{Y}_{N+6} + (-q_1+a)^{-1}(R_{N+4}-u_1) = \\ &= -q_2\dot{Y}_{N+6} + u_2 \end{aligned}$$

and generally

$$\dot{Y}_{N+2i} = -q_i\dot{Y}_{N+2i+2} + u_i; \qquad i \in [1,\ N/2], \tag{42}$$

where the unknown $\dot{Y}_{2N+2}$, corresponding to the largest value of the index i should be replaced by zero.

The recursive formulae, analogous to (38) and (39), are as follows:

$$\left.\begin{aligned} & q_1 = (-q_{i-1}+a)^{-1}; \qquad i \in [1,\ N/2-1], \\ & q_1 = a^{-1}; \\ & q_{N/2} = (-q_{N/2-1}+b)^{-1} \end{aligned}\right\} \tag{43}$$

and

$$\left.\begin{aligned} & u_i = q_i(R_{N+2i}-u_{i-1}); \qquad i \in [1,\ N/2-1], \\ & q_0 = 0. \end{aligned}\right\} \tag{44}$$

The remark concerning the possibility of storing the inverses (43) remains valid in this case, too.

Note that the systems (34), (35) can be solved parallelly.

3. The equation $\mathbf{B}\dot{\mathbf{Y}} = \mathbf{V}$

Matrix **B** in (13) is defined (*loc. cit.*) by means of the following three submatrices of the order N:

$$\beta_1^{[N]} \equiv \alpha_1{}^{[N]}, \tag{45}$$

where the right hand side is denoted by (19);

$$\beta_2^{[N]} = 4\beta_1^{[N]} \tag{46}$$

and

$$\alpha_2^{[N]} = \begin{bmatrix} 4 & 1 & & & & 0 \\ 1 & 4 & 1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & 1 & 4 & 1 \\ 0 & & & & 1 & 4 \end{bmatrix}. \tag{47}$$

By the use of these submatrices, system (13) of ordinary differential equations can be rewritten in the following manner:

$$\left.\begin{aligned}
&\beta_2\dot{Y}_1 + \beta_1\dot{Y}_2 + \alpha_2\dot{Y}_{N+2} = V_1;\\
&\beta_1\dot{Y}_1 + \beta_2\dot{Y}_2 + \beta_1 Y_3 - \alpha_2\dot{Y}_{N+1} + \alpha_2\dot{Y}_{N+3} = V_2;\\
&\beta_1\dot{Y}_2 + \beta_2\dot{Y}_3 + \beta_1 Y_4 - \alpha_2\dot{Y}_{N+2} + \alpha_2\dot{Y}_{N+4} = V_4;\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\beta_1\dot{Y}_{N-2} + \beta_2\dot{Y}_{N-1} + \beta_1\dot{Y}_N - \alpha_2\dot{Y}_{2N-2} + \alpha_2\dot{Y}_{2N} = V_{N-1};\\
&\beta_1\dot{Y}_{N-1} + \beta_2\dot{Y}_N - \alpha_2\dot{Y}_{2N-1} = V_N,
\end{aligned}\right\} \tag{48}$$

$$\left.\begin{aligned}
&- \alpha_2\dot{Y}_2 + \beta_2\dot{Y}_{N+1} + \beta_1\dot{Y}_{N+2} = V_{N+1};\\
&\alpha_2\dot{Y}_1 - \alpha_2\dot{Y}_3 + \beta_1 Y_{N+1} + \beta_2\dot{Y}_{N+2} + \beta_1\dot{Y}_{N+3} = V_{N+2};\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\alpha_2\dot{Y}_{N-2} - \alpha_2\dot{Y}_N + \beta_1 Y_{2N-2} + \beta_2\dot{Y}_{2N-1} + \beta_1\dot{Y}_{2N} = V_{2N-1};\\
&\alpha_2\dot{Y}_{N-1} + \beta_1 Y_{2N-1} + \beta_2\dot{Y}_{2N} = V_{2N}.
\end{aligned}\right\} \tag{49}$$

The matrix of this system is shown in Fig. 4 for the case $N = 8$. Again, the empty squares in this Figure ought in fact to be filled by null submatrices of the order N.

$$\mathbf{B}^{[8]} = \left[\begin{array}{cccccccccccccccc}
\beta_2 & \beta_1 & & & & & & & & \alpha_2 & & & & & & \\
\beta_1 & \beta_2 & \beta_1 & & & & & & -\alpha_2 & & \alpha_2 & & & & & \\
 & \beta_1 & \beta_2 & \beta_1 & & & & & & -\alpha_2 & & \alpha_2 & & & & \\
 & & \beta_1 & \beta_2 & \beta_1 & & & & & & -\alpha_2 & & \alpha_2 & & & \\
 & & & \beta_1 & \beta_2 & \beta_1 & & & & & & -\alpha_2 & & \alpha_2 & & \\
 & & & & \beta_1 & \beta_2 & \beta_1 & & & & & & -\alpha_2 & & \alpha_2 & \\
 & & & & & \beta_1 & \beta_2 & \beta_1 & & & & & & -\alpha_2 & & \alpha_2 \\
 & & & & & & \beta_1 & \beta_2 & & & & & & & -\alpha_2 & \\
 & -\alpha_2 & & & & & & & \beta_2 & \beta_1 & & & & & & \\
\alpha_2 & & -\alpha_2 & & & & & & \beta_1 & \beta_2 & \beta_1 & & & & & \\
 & \alpha_2 & & -\alpha_2 & & & & & & \beta_1 & \beta_2 & \beta_1 & & & & \\
 & & \alpha_2 & & -\alpha_2 & & & & & & \beta_1 & \beta_2 & \beta_1 & & & \\
 & & & \alpha_2 & & -\alpha_2 & & & & & & \beta_1 & \beta_2 & \beta_1 & & \\
 & & & & \alpha_2 & & -\alpha_2 & & & & & & \beta_1 & \beta_2 & \beta_1 & \\
 & & & & & \alpha_2 & & -\alpha_2 & & & & & & \beta_1 & \beta_2 & \beta_1 \\
 & & & & & & \alpha_2 & & & & & & & & \beta_1 & \beta_2
\end{array}\right]$$

Fig. 4. An example – for $N = 8$ – of the matrix of coefficients in (13), partitioned into submatrices of the order N

As in the previous Section, the symbols

$$\dot{Y}_k, \quad V_k; \qquad k \in [1,\, 2N]$$

denote column vectors, each one consisting of N real elements.

It will now be demonstrated that the unknowns

$$\dot{Y}_k; \qquad k \in [1, N] \tag{50}$$

can be eliminated from system (48)–(49), and a 'reduced' system will be derived containing solely the other half of the unknowns

$$\dot{Y}_k; \qquad k \in [N+1,\, 2N]. \tag{51}$$

For this purpose, subsystem (48) will be multiplied by the inverse β_1^{-1} of submatrix (45), and subsystem (49) by the inverse α_2^{-1} of submatrix (47). By introducing for the sake of convenience the following notation:

$$\left.\begin{aligned} &a = \beta_1^{-1}\alpha_2; &&b = \alpha_2^{-1}\beta_1;\\ &\beta_1^{-1}V_k = W_k; &&k \in [1,\, N];\\ &\alpha_2^{-1}V_k = W_k; &&k \in [N+1,\, 2N], \end{aligned}\right\} \tag{52}$$

and performing the said multiplications, one arrives easily at a new form of system (48)–(49):

$$\left.\begin{aligned} 4\dot{Y}_1 + \dot{Y}_2 \qquad\qquad + a\dot{Y}_{N+2} &= W_1;\\ \dot{Y}_1 + 4\dot{Y}_2 + \dot{Y}_3 - a\dot{Y}_{N+1} + a\dot{Y}_{N+3} &= W_2;\\ \dot{Y}_2 + 4\dot{Y}_3 + \dot{Y}_4 - a\dot{Y}_{N+2} + a\dot{Y}_{N+4} &= W_3;\\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\ \dot{Y}_{N-2} + 4\dot{Y}_{N-1} + \dot{Y}_N - a\dot{Y}_{2N-2} + a\dot{Y}_{2N} &= W_{N-1};\\ \dot{Y}_{N-1} + 4\dot{Y}_N \qquad - a\dot{Y}_{2N-1} \qquad &= W_N, \end{aligned}\right\} \tag{53}$$

$$\left.\begin{aligned} -\dot{Y}_2 \qquad\qquad + 4b\dot{Y}_{N+1} + b\dot{Y}_{N+2} &= W_{N+1};\\ \dot{Y}_1 - \dot{Y}_3 + b\dot{Y}_{N+1} + 4b\dot{Y}_{N+2} + b\dot{Y}_{N+3} &= W_{N+2};\\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\ \dot{Y}_{N-2} - \dot{Y}_N + b\dot{Y}_{2N-2} + 4b\dot{Y}_{2N-1} + b\dot{Y}_{2N} &= W_{2N-1};\\ \dot{Y}_{N-1} \qquad + b\dot{Y}_{2N-1} + 4b\dot{Y}_{2N} \qquad &= W_{2N}. \end{aligned}\right\} \tag{54}$$

It should be emphasised that the same symbol W_k has quite a different meaning in the present Section than in the previous one.

The next step in the elimination of unknowns (50) consists in solving subsystem (54) with respect to those unknowns. This can be done separately for the *even* and for the *odd* indices at these unknowns.

In the first case, the following general formulae can be derived:

$$\dot{Y}_k = bR_k + T_k; \qquad k = 2,\, 4, ..., N, \tag{55}$$

where

$$\left.\begin{aligned} &T_2 = -W_{N+1}, \\ &T_k = T_{k-2} - W_{N+k-1}; \qquad k = 4,\, 6, ...N, \end{aligned}\right\} \tag{56}$$

$$\left.\begin{aligned} &R_2 = 4\dot{Y}_{N+1} + \dot{Y}_{N+2}, \\ &R_k = (\dot{Y}_{N+k-2} + 4\dot{Y}_{N+k-1} + \dot{Y}_{N+k}) + R_{k-2}; \quad k = 4, 6, ..., N. \end{aligned}\right\} \tag{57}$$

Likewise, in the second case:

$$\dot{Y}_k = bR_k + T_k; \qquad k = N-1, \quad N-3, ..., 1, \tag{58}$$

where

$$\left.\begin{aligned} &T_{N-1} = W_{2N}, \\ &T_k = T_{k+2} + W_{N+k+1}; \qquad k = N-3, \quad N-5, ..., 1, \end{aligned}\right\} \tag{59}$$

$$\left.\begin{aligned} &R_{N-1} = -(\dot{Y}_{2N-1} + 4\dot{Y}_{2N}), \\ &R_k = -(\dot{Y}_{N+k} + 4\dot{Y}_{N+k+1} + \dot{Y}_{N+k+2}) + R_{k+2}; \quad k\ = N-3, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad N-5, ..., 1. \end{aligned}\right\} \tag{60}$$

Substitution of (55)–(57) and (58)–(60) into (53) yields the sought-after *reduced* system of equations, which does not contain unknowns (50). After multiplication by the inverse b^{-1} of submatrix b in (52) the reduced system can be presented in a relatively simple form. The right hand sides follow from the formulae

$$\left.\begin{aligned} &S_k = b^{-1}\Big[W_k - (T_{k-1} + 4T_k + T_{k+1})\Big]; \qquad k \in [1,\, N], \\ &T_o = T_{N+1} = 0 \end{aligned}\right\} \tag{61}$$

and the matrix of coefficients is shown in Fig. 5 for the case $N = 8$. The new submatrices in the Figure are defined as follows:

$$\left.\begin{aligned} &E = 15\, I - b^{-1}a, \\ &F = 3\, I - b^{-1}a. \end{aligned}\right\} \tag{62}$$

The main diagonal of the matrix $\mathbf{B}_R{}^{[8]}$ in Fig. 5 consists of null submatrices and it is true for any even number N. However, the determinant of such a matrix does not vanish:

$$\left|\mathbf{B}_R^{[N]}\right| \neq 0; \qquad N - \text{even}, \tag{63}$$

an analogous property being mentioned in the previous Section. In other words, the matrix is *nonsingular* for even values of N.

$$\mathbf{B}_R^{[8]} = \begin{bmatrix} 0 & -E & -8I & -16I & -8I & -16I & -8I & -16I \\ E & 0 & -F & -8I & -4I & -8I & -4I & -8I \\ 8I & F & 0 & -E & -8I & -16I & -8I & -16I \\ 16I & 8I & E & 0 & -F & -8I & -4I & -8I \\ 8I & 4I & 8I & F & 0 & -E & -8I & -16I \\ 16I & 8I & 16I & 8I & E & 0 & -F & -8I \\ 8I & 4I & 8I & 4I & 8I & F & 0 & -E \\ 16I & 8I & 16I & 8I & 16I & 8I & E & 0 \end{bmatrix}$$

Fig. 5. An example – for $N = 8$ – of the matrix $\mathbf{B}_R$ of coefficients of the reduced system of equations, partitioned into submatrices of the order N

Considering the matrix in Fig. 5 one comes easily to the conclusion that it can be partitioned in a more convenient way by the use of just five submatrices:

$$\left.\begin{aligned} &\alpha = \begin{bmatrix} 0 & -E \\ E & 0 \end{bmatrix}; &&\beta = \begin{bmatrix} 8I & F \\ 16I & 8I \end{bmatrix}; \\ &\gamma = \begin{bmatrix} 8I & 4I \\ 16I & 8I \end{bmatrix}; &&\delta = -\begin{bmatrix} 8I & 16I \\ F & 8I \end{bmatrix} = -\beta^T; \\ &\varepsilon = -\begin{bmatrix} 8I & 16I \\ 4I & 8I \end{bmatrix} = -\gamma^T && \end{aligned}\right\} \qquad (64)$$

of the order $2N$. It should be recalled that all the submatrices

$$E,\ F,\ I,\ 0 \qquad (65)$$

are of the order N. The new partitioning of the matrix $\mathbf{B}_R{}^{[8]}$ is shown in Fig. 6. It is clear that the main diagonal consists solely of the submatrices α. Similarly, the two closest diagonals contain the submatrices β and δ. The remaining space is filled by the submatrices γ and ε in the manner indicated in Fig. 6. The new partitioning does not influence the time necessary for the solution of the reduced system of equations. Nevertheless, construction of its matrix of coefficients becomes easy and systematic.

The matrix $\mathbf{B}_R^{[N]}$ does not appear to possess properties such as tridiagonality, enabling one to solve the reduced system of equations in a special time-saving manner, as was the case in the previous Section. Therefore, one of the existing methods of solution has to be applied. The one selected by ourselves involves representing the matrix of coefficients by means of the product of two triangular matrices, as is shown for the case $N = 8$ (Fig. 7).

$$\mathbf{B}_R{}^{[8]} = \begin{bmatrix} \alpha & \delta & \varepsilon & \varepsilon \\ \beta & \alpha & \delta & \varepsilon \\ \gamma & \beta & \alpha & \delta \\ \gamma & \gamma & \beta & \alpha \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{13} & B_{14} \\ B_{21} & B_{22} & B_{23} & B_{24} \\ B_{31} & B_{32} & B_{33} & B_{34} \\ B_{41} & B_{42} & B_{43} & B_{44} \end{bmatrix}$$

Fig. 6. An example – for $N = 8$ – of the matrix $\mathbf{B}_R$ of coefficients of the reduced system, partitioned into submatrices of the order $2N$

$$\mathbf{B}_R{}^{[8]} = \begin{bmatrix} L_{11} & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 \\ L_{31} & L_{32} & L_{33} & 0 \\ L_{41} & L_{42} & L_{43} & L_{44} \end{bmatrix} \begin{bmatrix} I & U_{12} & U_{13} & U_{14} \\ 0 & I & U_{23} & U_{24} \\ 0 & 0 & I & U_{34} \\ 0 & 0 & 0 & I \end{bmatrix}$$

Fig. 7. An example – for $N = 8$ – of the matrix $\mathbf{B}_R$ of coefficients of the reduced system, represented by product of two triangular matrices

It can be referred to as Banachiewicz's method – see Zurmühl (1949) and Prosnak (1993) – generalised, however, for systems of equations in which every term is represented by a column vector consisting of real elements.

Computation of the two factors of the product

$$\mathbf{B}_R^{[N]} = \mathbf{L}^{[N]}\,\mathbf{U}^{[N]} \tag{66}$$

can be treated as a routine problem of matrix algebra, and will be omitted. However, the way of dealing with the product (66) should be perhaps recalled.

For this purpose, the reduced system of equations will be rewritten in the following form:

$$\mathbf{LU\dot{Y}} = \mathbf{S}, \tag{67}$$

where

$$\left.\begin{aligned} \dot{\mathbf{Y}} &= [\widetilde{Y}_1,\, \widetilde{Y}_2, ...,\, \widetilde{Y}_{N/2}]^T; \\ \mathbf{S} &= [\widetilde{S}_1,\, \widetilde{S}_2, ...,\, \widetilde{S}_{N/2}]^T, \end{aligned}\right\} \tag{68}$$

the subvectors $\widetilde{Y}_k$, $\widetilde{S}_k$ containing $2N$ real elements, unlike subvectors (51), (61), which consist of N real elements only. This follows from the new partition of matrix $\mathbf{B}_R^{[N]}$, *cf.* Figs. 5 and 6. The superscripts $[N]$ at $\mathbf{L}$ and $\mathbf{U}$ in (67) are omitted for the sake of simplicity. Now, an auxiliary unknown

$$\mathbf{U}\dot{\mathbf{Y}} = \mathbf{X} \tag{69}$$

will be substituted into (67), which yields

$$\mathbf{L}\mathbf{X} = \mathbf{S}. \tag{70}$$

Remembering that $\mathbf{L}$ denotes the lower triangular matrix, one can see at once that determination of the solution of (70) with respect to $\mathbf{X}$ is very simple.

Substitution of this solution into (69) yields the equation with the unknown $\dot{\mathbf{Y}}$ which can also be computed quite easily by virtue of the properties of the matrix $\mathbf{U}$.

Finally, vectors (50) have to be evaluated by means of recursive formulae (55)–(57) and (58)–(60).

4. Numerical experiments and some conclusions

In order to obtain information on the algorithms presented in the previous Sections and, in particular, on the time necessary for their numerical realisation, measurement of time was included in the computer programs already mentioned. The results of these measurements are collected in Table 1.

Table 1. Time necessary to perform selected operations

N	matrix	t_1	t_2	t_3	t_4	t_5	t_6	t_7
8	**A**	0.010	1.883	0.027	0.233	0.034	–	–
	B	0.020	3.435	0.028	1.371	–	0.106	0.174
10	**A**	0.030	6.179	0.066	0.452	0.067	–	–
	B	0.030	11.667	0.066	4.317	–	0.227	0.716
20	**A**	0.440	476.916	1.105	8.524	0.701	–	–
	B	0.571	784.368	1.106	146.352	–	3.128	45.733

It should be recalled and emphasised that the left hand sides in (12), (13), (14) do not contain the kinematic viscosity ν. This means that the algorithms do not depend either on this constant or on the Reynolds Number

$$Re = \frac{V_s\, l}{\nu}, \tag{71}$$

where V_s denotes a velocity scale characterising the velocity distribution (9), and $l = 1$ the length of the side of the square in Fig. 1. In fact, the algorithms will work for any right hand sides in eqs. (12), (13), (14), even for those which have no connection at all with the original problem presented in Section 1. Therefore, the test computations which served as a basis for Table 1, were performed for column vectors **W**, **V** and **U** in the equations just cited, every vector containing $2N^2$ zeroes except one element equal to 1. The calculations were repeated $2N^2$ times, the non-vanishing element occupying correspondingly the row numbered 1, 2, 3, ..., $2N^2$. The total time of these calculations was subsequently divided by their number, and such mean times required for just one solution of the corresponding equation appear in Table 1.

The symbols **A**, **B** in the Table refer to the matrices in eqs. (12)–(13); the other symbols denote time in seconds:

t_1 – time required to generate the matrix;

t_2 – time required to compute the inverse of the matrix;

t_3 – time required to evaluate the solutions of the respective eqs. (12)–(13) by the use of the inverse, in accordance with (16)–(17);

t_4 – time required to solve eq. (12) or (13) by the use of the Gauss elimination method;

t_5 – time required to solve eq. (12) by means of the method presented in Section 2: the initial system of equations is reduced by half, and the matrix of coefficients of the system thus reduced is transformed to two tridiagonal ones;

t_6 – time required to solve eq. (13) by means of the method presented in Section 3: the initial system of equations is reduced by half, and the matrix of coefficients of the system thus reduced is replaced by the product of two triangular matrices;

t_7 – time required to solve eq. (13) by means of the method presented in Section 3; the reduced system of equations is solved directly by the use of the Gauss elimination without any preliminary transformation of the reduced system.

It follows from Table 1 that the shortest times necessary to solve eqs. (12)–(13) correspond to methods based on special properties of the respective matrices of coefficients, *i.e.* to the methods presented in Sections 2 and 3.

Comparison of the times necessary for a single solution of these equations by the use of one of these methods with the times needed just to invert the corresponding matrix gives some idea of the possible gain. For instance, when $N = 20$ and for the eq. (12), there is

$t_2/t_5 = 476.916/0.701 \cong 680.$

For eq. (13) and the same value of N this ratio is smaller:

$t_2/t_6 = 784.368/3.128 \cong 250.$

Taking into account the respective times necessary to evaluate the vector of unknowns after (16) and (17), *i.e.*

$t_3 = 1.105 \quad \text{and} \quad t_3 = 1.106,$

one can see at once, that such an evaluation is undoubtedly impractical in the first case, especially if one considers also further elements, such as the relatively small storage required by these two methods in comparison with the other ones.

Moreover, the corresponding numerical calculations apply only to matrices of the order N in the case of eq. (12), and of the order $2N$ in the case of eq. (13).

Hence, the *general conclusion* can be drawn that the both presented algorithms should be regarded as an achievement of the aim consisting in improvement of the solution to the test problem under consideration – with respect to economy both in computation time and in the storage required.

It is to be hoped that at least the underlying ideas can also be utilised in applications of the new method other than the driven cavity problem.

Some further practical conclusions can be drawn from Table 1. It should be noted that the widely applied method of Gauss elimination is rather time consuming, especially in comparison with the usual form of Banachiewicz's method. The same conclusion is true as far as computing the inverses is concerned, this operation being performed in our programs also by the use of Gauss elimination.

It should be emphasised that the *general assumption* underlying the considerations in the present paper concerns the independence of matrices **A** and **B** of time. The conclusions arrived at are valid for this assumption. They should be not extended to the case of non-stationary matrices, when *e.g.* the application of the method based on formulae (16)–(17) would be inadvisable. On the other hand, the two methods presented in Sections 2 and 3 would emphasise their merits even more.

Acknowledgements

Cordial thanks are due to Mrs Mirosława Kołodziejczyk PhD, MSc (Eng) for providing the authors with the results contained in Table 1. The difficult task accomplished by Mrs Krystyna Borkowska, *i.e* the preparation of the typescript, is highly appreciated too.

References

Klonowska M. E., Kołodziejczyk M., *Tests of a new method for the numerical solution of the Navier-Stokes equations. Part I: plane unsteady flow in finite, rectangular domains*, Appl. Mech. Eng., (in press).

Lambert J. D., 1973, *Computational methods in ordinary differential equations*, John Wiley & Sons, London–New York–Sydney, 278 pp.

Prosnak W. J., 1993, *Introduction to numerical fluid mechanics. Part A: fundamental numerical methods*, Ossolineum, Wrocław, 490 pp., (in Polish).

Prosnak W. J., Kosma Z. J., 1991, *On a new method for the numerical solution of the Navier-Stokes equations*, Acta Mech., 89, 45–63.

Zurmühl R., 1949, *Zur numerischen Auflösung linearer Gleichungssysteme nach dem Matrizenverfahren von Banachiewicz*, Z. angew. Math. Mech., 23, 76–84.