

# ON THE ECONOMICAL SOLUTION METHOD FOR A SYSTEM OF LINEAR ALGEBRAIC EQUATIONS

JAN AWREJCEWICZ, VADIM A. KRYSKO, AND ANTON V. KRYSKO

*Received 30 March 2004*

The present work proposes a novel optimal and exact method of solving large systems of linear algebraic equations. In the approach under consideration, the solution of a system of algebraic linear equations is found as a point of intersection of hyperplanes, which needs a minimal amount of computer operating storage. Two examples are given. In the first example, the boundary value problem for a three-dimensional stationary heat transfer equation in a parallelepiped in  $\mathbb{R}^3$  is considered, where boundary value problems of first, second, or third order, or their combinations, are taken into account. The governing differential equations are reduced to algebraic ones with the help of the finite element and boundary element methods for different meshes applied. The obtained results are compared with known analytical solutions. The second example concerns computation of a nonhomogeneous shallow physically and geometrically nonlinear shell subject to transversal uniformly distributed load. The partial differential equations are reduced to a system of nonlinear algebraic equations with the error of  $O(h_{x_1}^2 + h_{x_2}^2)$ . The linearization process is realized through either Newton method or differentiation with respect to a parameter. In consequence, the relations of the boundary condition variations along the shell side and the conditions for the solution matching are reported.

## 1. Introduction

It is obvious that a vast number of problems in physics, mechanics, and technology is modelled through linear and nonlinear partial differential equations (PDEs, equations of mathematical physics). In the next step, PDEs are usually reduced to a linear or nonlinear set of algebraic equations applying either finite element or finite difference methods. In what follows, only the method enabling solution of a large number of algebraic equations is considered. To solve the mentioned problem, there are two classical approaches that are direct and iterational methods. Perhaps the most popular are the Gauss reduction for a general system, and the relaxation or relaxation matrix methods usually applied for triple-diagonal or block-diagonal matrices. The mentioned methods belong to the direct ones. Their efficiency depends on the system equation order and on the matrix structure.

From the point of view of iterative methods, a system of equations is treated as the first-order operator equations of the form  $Ax = b$ , where  $A = (a_{ik})$  is a square matrix of dimension  $n \times n$ ,  $x = (x_1, x_2, \dots, x_n)$  stands for the sought vector, and  $b = (b_1, b_2, \dots, b_n)$  is a given vector (right-hand side). In the case of the mentioned iterative methods, the system of algebraic equations can be treated as the first-order operator equations with the operator acting in an  $n$ -dimensional space  $H_n(A : H_n \rightarrow H_n)$ ,  $x, b \in H_n$ .

It is worth noticing that the application of the general theory makes it possible to prove convergence of iterations for Seidel's and upper relaxation methods with the minimal constraints used by operator  $A$ . Usually, two types of approaches are applied: (i) for known boundaries  $\gamma_1 > 0$  and  $\gamma_1 \geq \gamma_2$  for a spectrum of operator  $A$  lying in a certain energy space  $H_p$  and (ii) for the case when boundaries  $\gamma_1$  and  $\gamma_2$  are unknown. However, the triangle variational method seems to be more effective in application.

For all numerical methods, the reduction of the infinite-dimensional problem to that of the finite dimension plays a key role. It is expressed by the fact that a computational algorithm should yield a solution of the initial problem with a given accuracy  $\varepsilon > 0$  through a finite number  $Q(k)$  of actions. Moreover, the algorithm should be practically realized, that is, should yield a solution to the problem within the required computer time. Furthermore, the number of actions (and hence the time of solution)  $Q(\varepsilon)$  should be minimal for the considered problem. Algorithms with the mentioned properties are called economical. It is obvious that the choice of a numerical method of solving a system of linear algebraic equations depends on many circumstances, that is, on the matrix  $A$  properties, on the computational type applied, and so forth.

A computational type is one of the possible formulations of the problem, like finding a solution to one special problem of  $Ax = b$ , or finding solutions to a few variants of problem  $Ax = b$  with the same matrix  $A$  and different right-hand sides  $B$ . It may happen that a nonoptimal choice of the problem with one matrix  $Ax = b$  can be suitable for multivariant computations.

Note that for multivariant computations, one may decrease the average number of operations for one variant if some quantities are conserved and not computed once more for each variant. It depends on a computer and its operating storage.

The choice of a computational algorithm should depend on the computational type, on the volume of the operating storage, and on the considered system order.

In this work, the presented method of solution of large algebraic equation systems offers reduction of the infinite-dimensional problem of mathematical physics to the finite-dimensional one through the finite difference approximation.

The used approach yields a band matrix possessing only a few diagonals with nonzero elements. Direct methods of solving those systems matched with the transformation of input matrix  $A$  do not enable arrival at similar results, that is, the possibility of solution of the system of high-order algebraic equations offered by them is rather doubtful.

## 2. Elimination method for equations

Consider a direct method [6], further referred to as the elimination method for equations, making it possible to find the system solution as a point of intersection of hyperplanes and

requiring a minimal operating storage volume. The latter property is rarely considered by the standard direct methods [3].

The idea of the proposed method will be illustrated and clarified when applied to a system of linear algebraic equations (SLAE).

Consider first SLAE I of the general form

$$\sum_{k=1}^n a_{ik}x_k = b_i \quad i = 1, 2, \dots, n. \tag{2.1}$$

The process of the proposed elimination consists of the following few steps.

*Step 1.* Assume  $j = n$ . Solve the  $j$ th equation of system (2.1)  $j$  times with respect to  $x_n$  by keeping variables  $x_1, x_2, \dots, x_{n-i}$  fixed. As a result, the set  $M_j$  consisting of points  $\{X^k\}$ ,  $k = k = 1, 2, \dots, j$ , with the coordinates

$$\begin{aligned} X^1(c, c, \dots, c, x_n^1), \\ X^2(p, c, \dots, c, x_n^2), \\ X^3(c, p, \dots, c, x_n^3), \\ \vdots \\ X^j(c, c, \dots, c, p, x_n^j), \end{aligned} \tag{2.2}$$

is obtained, where  $c, p$  are arbitrary numbers (it is suitable to take  $c = 1, p = -1$ ).

*Step 2.* Compute  $i = j - 1$  and construct set  $M$  of points that are intersections of straight lines going through points  $\{X^j, X^k\}$ ,  $k = 1, 2, \dots, i$ , and the hyperplane defined by the  $i$ th system of algebraic equations of the form

$$\begin{aligned} \sum_{l=1}^n a_{il}x_l = b_i, \\ x_l = x_l^j + (x_l^k - x_l^j)\lambda_k, \quad l = 1, 2, \dots, n, \end{aligned} \tag{2.3}$$

is solved. System (2.3) yields

$$\lambda_k = \frac{b_i - \sum_{l=1}^n a_{il}x_l^j}{\sum_{l=1}^n a_{il}(x_l^k - x_l^j)} \tag{2.4}$$

and the coordinates of the  $k$ th point of set  $M_i$  are computed in the following way:

$$\bar{x}_l^k = x_l^j + (x_l^k - x_l^j)\lambda_k, \quad l = 1, 2, \dots, n. \tag{2.5}$$

The set  $M_i$  of points  $\{X^k\}$ ,  $k = 1, 2, \dots, i$ , belongs to the hyperplane defined through the  $i$ th equation of system (2.1). The coordinates of the points of the mentioned set satisfy the last  $n - i$  equations of system (2.1). In Table 2.1, coordinates of set  $M_i$  for  $i = n - 1$  are reported.

Table 2.1. Coordinates of points in set  $M_i$  (for  $i = n - 1$ ).

	$X^1$	$X^2$	$X^3$	$\dots$	$X^{i-1}$	$X^i$
$x_1$	$c$	$x_1^2$	$c$	$\dots$	$c$	$c$
$x_2$	$c$	$c$	$x_2^3$	$\dots$	$c$	$c$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{n-3}$	$c$	$c$	$c$	$\dots$	$x_{n-3}^{i-1}$	$c$
$x_{n-2}$	$c$	$c$	$c$	$\dots$	$c$	$x_{n-2}^i$
$x_{n-1}$	$x_{n-1}^1$	$x_{n-1}^2$	$x_{n-1}^3$	$\dots$	$x_{n-1}^{i-1}$	$x_{n-1}^i$
$x_n$	$x_n^1$	$x_n^2$	$x_n^3$	$\dots$	$x_n^{i-1}$	$x_n^i$

Step 3. Assume  $j = i$  and go to Step 2.

Repeating computations  $n - 1$  times, one gets a point in an  $n$ -dimensional space, which is a solution to system (2.1).

The above considerations allow the conclusion that the above algorithm is suitable for solving SLAE through computations of the matrix rows.

Observe also that for a chosen parametrization of set  $M_n$  and a given method applied for the construction of the straight lines, only the coordinates  $x_{k-1}^k, x_i^k, x_{i+1}^k, \dots, x_n^k$  should be determined; other coordinates preserve values equal to  $c$ .

It is worth noticing that the proposed novel approach becomes four times smaller, which enables a considerable decrease in the volume of the computer operating storage used to keep coordinates of points  $X^k$ .

Consider SLAE II with a band matrix:

$$\sum_{s=1}^l a_{jj+i, x_{j+i}} = b_j, \quad j = 1, 2, \dots, n, \quad 1 \leq j + i_s \leq n, \quad i_s \in I, \quad (2.6)$$

where  $l$  is the number of nonzero diagonals and  $I$  is a set consisting of  $l$  elements, that is, ordered numbers of nonzero band diagonals with an account of sign with respect to the main diagonal.

The following parameters are further introduced:  $l_1, m_1$  are the numbers of nonzero diagonals and width of the band part lying under the main diagonal, respectively;  $l_2, m_2$  are the numbers of nonzero diagonals and width of the band part lying over the main diagonal; and  $m = m_1 + m_2 + 1$  is the bandwidth.

Observe that

$$m_1 = -I(1), \quad m_2 = I(l). \quad (2.7)$$

The algorithm for solving SLAE (2.6) runs as follows.

Step 1. Assume  $j = n$  and solve the  $j$ th equation of (2.6)  $m_1 + 1$  times with respect to  $X_n$  introducing variables  $x_{n-m_1}, x_{n-m_1+1}, \dots, x_{n-1}$ . As a result, the set  $M_j$  of points  $X^k$  in

the  $(m_1 + 1)$ th-dimensional space of variables  $x_{n-m_1}, x_{n-m_1+1}, \dots, x_n$  is obtained:

$$\begin{aligned} & X^{j-m_1}(c, c, \dots, c, c, x_n^{j-m_1}), \\ & X^{j-m_1+1}(p, c, \dots, c, c, x_n^{j-m_1+1}), \\ & \quad \vdots \\ & X^{j-1}(c, c, \dots, p, c, x_n^{j-1}), \\ & X^j(c, c, \dots, c, p, x_n^j). \end{aligned} \tag{2.8}$$

Note that it is advisable to take  $c = 1, p = -1$ .

Step 2. Compute  $i = j - 1$ . If  $i > m_1$ , then the set  $M_j$  is modified by introducing the variable

$$x_{j-m_1-1} = c \tag{2.9}$$

and the points

$$\begin{aligned} & X^{j-m_1-1} = X^{j-m_1}, \\ & X^{j-m_1}(p, c, c, \dots, c, c, x_n^{j-m_1}). \end{aligned} \tag{2.10}$$

A new set  $\tilde{M}_j$  of points  $\{X^k\}$ ,  $k = j - m_1 - 1, j - m_1, \dots, j$ , is obtained in the space of variables  $x_{n-m_1-1}, x_{n-m_1}, \dots, x_n$  of the form

$$\begin{aligned} & X^{j-m_1-1}(c, c, c, \dots, c, c, x_n^{j-m_1-1}), \\ & X^{j-m_1}(p, c, c, \dots, c, c, x_n^{j-m_1}), \\ & \quad \vdots \\ & X^{j-1}(c, c, c, \dots, p, c, x_n^{j-1}), \\ & X^j(c, c, c, \dots, c, p, x_n^j). \end{aligned} \tag{2.11}$$

Construct a set of points attained through intersection of the hyperplane defined by the  $(j - m)$ th equation of system (2.6) with the straight lines going through points  $\{X^j, X^k\}$ ,  $k = j - m_1 - 1, j - m_1, \dots, j - 1$ . For this purpose,  $i$ , or  $m_1 + 1$  if  $i > m_1$ , systems of equations

$$\begin{aligned} & \sum_{s=1}^l a_{ii+i_s} x_{i+i_s} = b_i, \\ & x_{i+i_s} = x_{i+i_s}^j + (x_{i+i_s}^k - x_{i+i_s}^j) \lambda_k, \quad s = 1, 2, \dots, l, \quad 1 \leq i + i_s \leq n, \end{aligned} \tag{2.12}$$

are solved. System (2.12) yields

$$\lambda_k = \frac{b_i - \sum_{l=1}^n a_{ii+i_s} x_{i+i_s}^j}{\sum_{l=1}^n a_{ii+i_s} (x_{i+i_s}^k - x_{i+i_s}^j)}. \tag{2.13}$$

Table 2.2. Coordinates of points  $M_j$  (for  $m_1 < j \ll n - m_2$ ).

	$X^{k-1}$	$X^k$	$X^{k+1}$	$\dots$	$X^{j-1}$	$X^j$
$x_{k-1}$	$c$	$p$	$c$	$\dots$	$c$	$c$
$x_k$	$c$	$c$	$x_k^{k+1}$	$\dots$	$c$	$c$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{j-2}$	$c$	$c$	$c$	$\dots$	$x_{j-2}^{j-1}$	$c$
$x_{j-1}$	$c$	$c$	$c$	$\dots$	$c$	$x_{j-1}^j$
$x_j$	$x_j^{k-1}$	$x_j^k$	$x_j^{k+1}$	$\dots$	$x_j^{j-1}$	$x_j^j$
$x_{j+1}$	$x_{j+1}^{k-1}$	$x_{j+1}^k$	$x_{j+1}^{k+1}$	$\dots$	$x_{j+1}^{j-1}$	$x_{j+1}^j$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{m-1}$	$x_{m-1}^{k-1}$	$x_{m-1}^k$	$x_{m-1}^{k+1}$	$\dots$	$x_{m-1}^{j-1}$	$x_{m-1}^j$
$x_m$	$x_m^{k-1}$	$x_m^k$	$x_m^{k+1}$	$\dots$	$x_m^{j-1}$	$x_m^j$

Then coordinates of the  $k$ th intersection point are computed:

$$\tilde{x}_r^k = x_r^j + (x_r^k - x_r^j)\lambda_k, \quad r = k, i, i + 1, \dots, i + m_2, r \leq n. \tag{2.14}$$

Step 3. Assume  $j = i$  and return to Step 2.

In Table 2.2, coordinates of the points of set  $M_j$  for  $m_1 < j \ll n - m_2$  are reported. Notice that for the repeated index combination, the following notation is introduced:  $k = j - m_1, m = j + m_2$ .

Remark. The already mentioned advantages with respect to the chosen parametrization of the input points and straight line construction methods hold also for this case.

Step 4. By repeating the computations through Steps 2 and 3  $n - 1$  times, a point in the space with coordinates  $x_1, x_2, \dots, x_{m_2}$  is obtained. The space with obtained coordinates overlaps with the values of the unknown  $x_1, x_2, \dots, x_{m_2}$  occurring in the solution of system (2.6).

Step 5. The order of the initial system is decreased first by neglecting  $m_2$  equations and by putting the found values of the  $m_2$  unknowns into the remaining equations. Notice that during this operation, the parameters of the truncated band matrix do not change.

Step 6. Parameters of the next group are chosen from the  $m_2$  unknowns when the computation through Steps 1, 2, and 3 of the truncated system is carried out.

The described procedure of the successive decrease in the system order is repeated until all values of  $n$  unknowns are found. Generally speaking, one may require to get  $n$  and  $m_2$  as integers. Notice that by adding a zero diagonal to the compressed band matrix, parameter  $m_2$  may be introduced arbitrarily.

Using the presented method (see [6]), it is possible to apply other variants of SLAE solution to the band matrix as well.

The proposed novel algorithm is stable with respect to the errors introduced by roundings. In this sense, as practical computations show, this method is close to that of

the Gauss elimination technique for unknowns with a partial choice of the main element. For instance, solving the system of equations considered in [6, page 61], for the matrix conditioned by  $\omega = 4.7 \cdot 10^5$ , the proposed method gives exact decimal digits, that is, of one order less than the Gauss method. This result is obtained owing to a particular symmetry of the fundamental computational formulas (2.13), (2.14) and a homogeneity in computations of all unknowns.

Owing to the algorithm, one may conclude that in order to store coordinates of the points of set  $M_j$ , one may isolate independently of the order of the sought system, that is, the set consisting of  $(m_1 + 2)(m_2 + 3)$  elements. From this point of view, the considered direct method is similar to iterative methods owing to the required volume of operational storage.

Furthermore, the proposed method is characterized by a real computational process cycling and a weak coupling of the algorithm due to constant and full refreshments of transitional results.

Indeed, giving the corresponding coordinates of the points of manifold  $M$  for a certain  $j < n$ , one may begin the process of solution of the system of equations from the equation with the number  $i = j - 1$ .

The mentioned property accounts for an essential increase in the solution process on a computer owing to the elimination of computation repetitions for truncated systems. Namely, while solving an input system of order  $n$ , we remember the coordinates of the points of the sets  $M_j$ ,  $j = n + 1 - lm_2$ ,  $l = 1, 2, \dots, n/m_2 - 2$ ,  $j \geq m_2 + 1Q$ .

In order to store the values of coordinates, it is advisable to use external computer memory. Next, after a successive truncated system of order  $n - k_{m_2}$ ,  $k = 1, 2, \dots$ , has been formulated, the corresponding set  $M_j$  of points  $j = 2m_2 + 1 + (k - 1)m_2$  is created.

The discussed algorithm for solving SLAE with a band matrix is also suitable for solving boundary value problems using the method of finite differences in a rectangular parallelepiped since the matrix of SLAE possesses a band with a regular structure. The latter property allows the use of one and only one set  $I$  for all equations.

For the space with a solution to a boundary value problem slightly differing from the canonical one, an artificial way of introducing fictitious equations for the mesh nodes to SLAE is as follows:

$$x_i = 0, \tag{2.15}$$

which supplements the given space to a canonical one.

We dwell now for a while on the computation of elements of set  $I$  ordered by the numbers of nonzero diagonals of a band of the SLAE matrix with respect to its main diagonal. In what follows, a three-dimensional case is considered.

Let  $n_1, n_2, n_3$  be the numbers of mesh nodes in directions  $ox_1, ox_2, ox_3$ , respectively. For approximation of partial derivatives of second order in the node  $\{i, j, k\}$ , the seventh point pattern is used:  $\{i, j, k; i \pm 1, j, k; i, j \pm 1, k; i, j, k \pm 1\}$ , where  $1 \leq i \leq n_1, 1 \leq j \leq n_2, 1 \leq k \leq n_3$ . Observe that the number of nonzero band diagonals is equal to that of nodes in the pattern  $l = 7$ . The position of nonzero diagonals with respect to the main diagonal of the band matrix is defined using ordered numbers of the unknowns corresponding to the nodes appearing in the pattern.

Numbering of unknowns is defined by the number of mesh nodes during SLAE formulation. Although various variants are possible, only those which formulate a band matrix of a regular structure are interesting, that is, one set  $I$  gives the position of nonzero coefficients in all equations.

The variant associated with the successive picking of nodes in directions  $ox_1, ox_2, ox_3$  and corresponding to the formula for the computation of unknown number  $n$  in the node  $\{i, j, k\}$  reads

$$s_{i,j,k} = i + (j - 1)n_1 + (k - 1)n_1n_2. \quad (2.16)$$

Formula (2.16) yields

$$\begin{aligned} i_1 &= s_{i,j,k-1} - s_{i,j,k} = -n_1n_2, \\ i_2 &= s_{i,j-1,k} - s_{i,j,k} = -n_1, \\ i_3 &= s_{i-1,j,k} - s_{i,j,k} = -1, \\ i_4 &= s_{i,j,k} - s_{i,j,k} = 0, \\ i_5 &= s_{i+1,j,k} - s_{i,j,k} = 1, \\ i_6 &= s_{i,j+1,k} - s_{i,j,k} = n_1, \\ i_7 &= s_{i,j,k+1} - s_{i,j,k} = n_1n_2. \end{aligned} \quad (2.17)$$

For the last band matrix parameters, one gets

$$m_1 = n_1n_2, \quad m_2 = m_1, \quad m = 2n_1(n_2 + 1), \quad n = n_1n_2n_3. \quad (2.18)$$

The formula used for the successive picking of nodes along directions  $ox_3, ox_1, ox_2$ , that is, the formula applied to compute an unknown number in the node  $\{k, i, j\}$ , is

$$s_{k,i,j} = j + (k - 1)n_3 + (j - 1)n_1n_3, \quad (2.19)$$

where elements  $i_s$  of set  $I$  read

$$\begin{aligned} i_1 &= -n_1n_3, & i_2 &= -n_3, & i_3 &= -1, \\ i_4 &= 0, & i_5 &= 1, & i_6 &= n_3, & i_7 &= n_2n_3, \end{aligned} \quad (2.20)$$

and the band parameter matrix follows:

$$m_1 = n_1n_3, \quad m_2 = m_1, \quad m = 2n_1n_3 + 1. \quad (2.21)$$

Finally, the last variant of the successive node picking in directions  $ox_2, ox_3, ox_1$  corresponds to the formula for computation of the unknown number in the node  $\{j, k, i\}$ , that is,

$$s_{j,k,i} = j + (k - 1)n_2 + (i - 1)n_2n_3, \quad (2.22)$$

where elements  $i_s$  of set  $I$  read

$$\begin{aligned} i_1 &= -n_2n_3, & i_2 &= -n_2, & i_3 &= -1, \\ i_4 &= 0, & i_5 &= 1, & i_6 &= n_2, & i_7 &= n_2n_3, \end{aligned} \quad (2.23)$$



Table 2.3. Exchange of mesh nodes.

Pattern	Row	Variant of the nodes exchange		
		$i, j, k$	$k, i, j$	$j, k, i$
Node	Element			
$i, j, k - 1$	1	1	3	2
$i, j - 1, k$	2	2	1	3
$i - 1, j, k$	3	3	2	1
$i, j, k$	4	4	4	4
$i + 1, j, k$	5	5	6	7
$i, j + 1, k$	6	6	7	5
$i, j, k + 1$	7	7	5	6

and the band element matrix has the form

$$m_1 = n_2 n_3, \quad m_2 = m_1, \quad m = 2n_2 n_3 + 1. \quad (2.24)$$

It follows from the algorithm description that both the parameter  $m$ -band SLAE matrix width and the system of order  $n$  determine the dimension of the operating storage and the machine time required to solve the equations on a computer.

Therefore, if values  $n_1, n_2, n_3$  are different, then the variant of mesh nodes preserving the minimal matrix bandwidth should be chosen. It is clear that the last picking procedure should be associated with the maximal number of nodes in direction  $n_x$ .

In practice, when the SLAE matrix in the form of a compressed band is built (only nonzero elements are considered), only one picking variant of the shell nodes (independently of the values of  $n_1, n_2, n_3$ ) is taken. Therefore, minimization of the bandwidth is introduced, which is reduced to the exchange of elements in rows, exchange of rows, and exchange of column elements of free terms.

In the considered example, the correspondence between elements of the band row is given in Table 2.3 for three variants of picking of mesh nodes. Relations (2.16), (2.19), and (2.22) are used for the exchange of rows and of column elements of free terms, which define simultaneously the ordering numbers of the system equations for the various variants of picking of the shell nodes.

If one considers the boundary value problem for a system of differential equations with respect to a few sought functions, then with the use of finite difference method, the structure of the SLAE matrix becomes dependent on the choice of the numbering of unknowns.

In the case of successive numeration of unknowns into groups corresponding to sought functions, the matrix has a block structure. The bandwidth exceeds the system order, which makes the considered algorithm noneffective.

Note that numeration of similar unknowns with a step equal to the number of sought functions corresponds to successive numeration of all (related to a mesh node) unknowns, and the SLAE matrix possesses a regular band-type structure. The bandwidth is essentially smaller than the system order, and hence the proposed method can be applied.

In practice, both numeration methods mentioned above are applied. The first one is used in the computation of nonzero elements of the equation matrix. The second one is applied during transformation of the matrix to a form suitable for the application of the direct solution method.

It is worth noticing that using the described approach, the general properties of the matrix transformation algorithm are applied independently of the number of the sought functions. The procedure consists of the following steps.

*Step 1.* Elements of set  $I$ , that is, ordering numbers of nonzero diagonals of the matrix with respect to the main diagonals, are computed. For this purpose (analogously to (2.16), (2.17)), ordering numbers of the unknowns corresponding to the nodes of the difference pattern scheme are computed first with respect to each group of the same unknowns in accordance with the first numbering order.

*Step 2.* Elements of set  $I_k$  are computed, that is, the relative numbers of these unknowns for the cases when the unknowns lie on the main diagonal associated with the first, the second, ..., the  $k$ th group of the unknowns named the same.

*Step 3.* A nonorder set  $I$  matching all sets  $I_k$  is constructed. Further, an increasing ordering of the elements of set  $I$  is carried out and the number of elements of set  $I$  with parameters  $l_1, l_2, m_1, m_2, m$  is computed.

*Step 4.* The reconstruction of the initial block-band compressed matrix is initiated through changes of row elements, rows, and free column elements with respect to the second numbering way for unknowns: first elements of all blocks, second elements of all blocks, and so on; first rows of blocks corresponding to the equations of the boundary value problem, second rows of these blocks, and so forth.

Elements of the column of free terms are transformed in a way similar to that applied to the rows.

*Step 5.* Then a band-type matrix in the compressed form is built in the following way: ordering numbers of the row elements from the block of the input matrix are defined through the condition of equality of the corresponding elements of sets  $I_k, I$ .

Notice that in its essential part, the resulting band is filled with zero elements.

Consider an example of the SLAE matrix transformation occurring during the solving of the boundary value problem for the system of two differential equations in partial derivatives of fourth order with variable coefficients in two-dimensional space using the method of finite differences. Denote by  $n_1, n_2$  the numbers of mesh nodes in directions  $ox_1, ox_2$ , respectively.

The derivatives in mesh node  $\{i, j\}$  are approximated through difference relations on the pattern consisting of 13 nodes (see Figure 2.1).

The following coordinates of the local nodes  $\{i_s, j_s\}, s = 1, 2, \dots, 13$ , are introduced:

$$\begin{aligned}
 i_1 &= i - 2, j_1 = j; i_2 = i - 1, j_2 = j - 1; i_3 = i - 1, j_3 = j; \\
 i_4 &= i - 1, j_4 = j + 1; i_5 = i, j_5 = j - 2; i_6 = i, j_6 = j - 1; \\
 i_7 &= i, j_7 = j; i_8 = i, j_8 = j + 1; i_9 = i, j_9 = j + 2; \\
 i_{10} &= i + 1, j_{10} = j - 1; i_{11} = i + 1, j_{11} = j; \\
 i_{12} &= i + 1, j_{12} = j + 1; i_{13} = i + 2, j_{13} = j.
 \end{aligned} \tag{2.25}$$

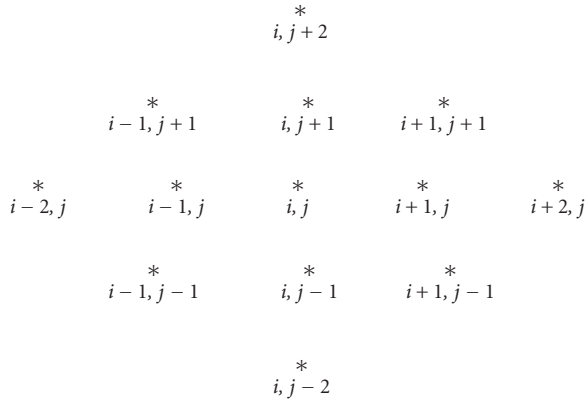


Figure 2.1. The pattern of 13 nodes.

The input matrix in the compressed form (nonzero diagonals) is built owing to the first numbering way for unknowns, and it describes a two-dimensional packet of dimension  $k_n \cdot k_l$ , where  $k = 2$  is the number of sought functions,  $n = n_1 n_2$  is the number of mesh nodes for which equations are constructed, and  $l = 13$  is the pattern dimension.

The discussed matrix is a block-diagonal one consisting of  $k^2$  blocks and possessing 39 nonzero diagonals.

Local index  $s$  of the pattern node defines a position in the row of a coefficient standing by the corresponding unknown in the equation for node  $\{i, j\}$ . The computation is carried out with respect to the beginning of the corresponding block.

Then, elements of the following sets are computed:

$$i_{1,k} \in I_1, \quad i_{2,k} \in I_2, \quad k = 1, 2, \dots, 26. \tag{2.26}$$

For this purpose, the relative members of all unknowns corresponding to pattern nodes and appearing in equations for node  $\{i, j\}$  are computed:

$$i_{1,k} = l_k - l_{13}, \quad i_{2,k} = l_k - l_{14}, \tag{2.27}$$

where even  $k : l_k = 2((i_s - 1)n_2 + j_s)$ , odd  $k : l_{k-1} = 2((i_s - 1)n_2 + j_s) - 1, s = 1, 2, \dots, 13$ .

The set  $I = I_1 \cup I_2$  is constructed and its elements are ordered in an increasing manner. The corresponding results for  $n_2 = 5$  are reported in the first four columns of Table 2.4.

As seen from this table, the number of nonzero diagonals of the band matrix is equal to 31. This value does not depend on  $n_2$  but is defined only through a type of the pattern chosen for approximation of derivatives and by the number of the functions sought in the boundary value problem.

Comparing elements of sets  $I_1, I_2$ , and  $I$ , one obtains a rule of distribution of the row elements in the modified matrix during construction of the band matrix in the compact form.

The correspondence between the ordering numbers of nonzero coefficients for the even and odd rows of the modified matrix and rows of the band matrix is reported in the

Table 2.4. Comparison of obtained results (see text).

$N$	$I_1$	$I_2$	$I$	$k = 1$	$k = 2$
1	-20	-21	-21	2	1
2	-19	-20	-20	3	2
3	-12	-13	-19	5	4
4	-11	-12	-13	6	5
5	-10	-1	-12	7	6
6	-9	-10	-11	8	7
7	-8	-9	-10	9	8
8	-7	-8	-9	10	9
9	-4	-5	-8	12	11
10	-3	-4	-7	13	12
11	-2	-3	-5	14	13
12	-1	-2	-4	15	14
13	0	-1	-3	16	15
14	1	0	-2	17	16
15	2	1	-1	18	17
16	3	2	0	19	18
17	4	3	1	20	19
18	5	4	2	21	20
19	8	7	3	23	22
20	9	8	4	24	23
21	10	9	5	25	24
22	11	10	7	26	25
23	12	11	8	27	26
24	13	12	9	28	27
25	20	19	10	30	29
26	21	20	11	31	30
27	—	—	12	—	—
28	—	—	13	—	—
29	—	—	19	—	—
30	—	—	20	—	—
31	—	—	21	—	—

last two columns of Table 2.4. The free positions in the row of the matrix are filled out by zeros. Notice that for arbitrary  $n_1, n_2$ , this correspondence is the same.

As known, when solving numerically boundary value problems for PDEs using finite element methods, there appears SLAE associated with matrices including a relatively small number of nonzero elements, most of which remain on the main diagonal, that

is, in this case, the equations dealt with are associated with the band matrix of irregular structure (the so-called cutting-off matrix).

However, there exists a modification of the elimination method for equations due to which it is possible to solve SLAE of a similar form. The cutting-off matrix will be introduced through two sets: the set of nonzero matrix elements and the corresponding set of indices, that is, relative numbers of nonzero elements with respect to the elements lying on the main diagonal, which are associated with the indices.

In what follows, contrary to the band matrix of a regular structure requiring only one row of indices for nonzero elements (set  $I$ ), in our case, this type of row is formulated for each of the equations (set  $I_i$ ). Here lies the fundamental difference in the application of the elimination method for the equations of SLAE using the cutting-off matrix.

We comment on parameters  $m_1$ ,  $m_2$ . Being the same for all equations, they are computed initially as

$$m_1 = \max(-I_i(1)), \quad m_2 = \max(I_i(l_i)), \quad (2.28)$$

where  $l_i$  denotes the number of nonzero coefficients in the  $i$ th equation. Therefore, the row of nonzero coefficients and the row of indices for each equation are supplemented, if necessary, by zeros in the case of elements, and by

$$I_i(1) = -m_1, \quad I_i(l_i) = m_2 \quad (2.29)$$

in the case of indices, where  $l_i$  is the value of  $I_i$  satisfying the above requirement.

As a result, SLAE with a band matrix of nonregular structure is obtained.

Note that the formulation of the cutting-off matrix with the help of two packets is optimal with respect to the memory storage volume since for the index conservation, it is sufficient that the machine word be 2 bytes long.

The corresponding computation shows, for instance, that for the conservation of a band-type matrix of the 27th order and a regular structure obtained through approximation of the Laplace operator in a three-dimensional space by finite difference equations, 5832 bytes are required in the general case, 1512 bytes are required in the case of its compressed form (only nonzero elements are considered), and finally 1350 bytes are needed when the matrix is represented by two packets.

It is worth noticing that the method of elimination extends essentially the possibility of computer computation needed to solve large-order SLAE through reaching a match between the advantages of both direct (universality, finiteness of the computational process, exactness) and iterative (minimal requirement of the storage volume) methods.

We finally emphasize particularities of the SLAE matrix in each separate case, which can also be treated as an advantage of the method.

### 3. Numerical solution of a three-dimensional equation of elliptic type

Many stationary processes of different physical properties lead to PDEs of elliptic type, to mention only the stationary problem of current distribution in a medium, problems of electrostatics and magnetostatics of the theory of plates and shells, problems of the theory of elasticity or the theory of filtration, and so forth.

Exact solution to the boundary value problems for elliptic equations can be found in rare cases only. Therefore, numerical methods are usually applied to solve differential equations of elliptic type (linear or nonlinear). In the latter case, equations are linearized through differentiation along the parameter [7] or using the “setup” method [2]. Hence, the problem is finally reduced to the solution of the SLAE.

Moreover, a majority of methods for nonlinear problem solving are reduced to a series of linear systems. Application of the method of finite differences to elliptic equations results in SLAE with a band-type matrix. In the case of PDEs, the matrix (SLAE) possesses the cutting-off band, with only a few nonzero elements (see Section 2).

We briefly model the problem of stationary heat distribution in a certain volume  $G$  with surface  $\Gamma$  of three-dimensional space  $x = (x_1, x_2, x_3)$ . The heat transfer equation is governed by the Fourier principle. A vector of heat stream density  $W$  is proportional to the temperature gradient  $V = V(x)$  such that

$$W = -K \operatorname{grad} V, \quad (3.1)$$

where  $K = K(x)$  is the heat transfer coefficient. The density of the heat stream is equal to the amount of the heat stream, a unit of time passing through a unit area of an isothermic surface.

To derive the equation governing the heat balance for certain volume  $U \in G$  and surface  $S$ , let the distributed heat sources exist inside volume  $U$  with density  $\varphi(x)$ , where  $\varphi(x)dU$  is the heat amount occurring in volume  $dU$ .

Let  $W_n$  describe the vector  $W$  projection onto external normal  $n$  to surface  $S$ . The heat balance equation is governed by the known rule, that is, the total heat stream passing through surface  $S$ ,

$$\iint_S W_n dS, \quad (3.2)$$

should be equal to the heat amount of

$$\iiint_U \varphi(x) dU \quad (3.3)$$

appearing in volume  $U$ , that is,

$$\iint_S W_n dS = \iiint_U \varphi(x) dU. \quad (3.4)$$

Using the Gauss formula  $\iint_S W_n dS = \iiint_U \operatorname{div} W$ , balance equation (3.4) is transformed to the form

$$\iiint_U (\operatorname{div} W - \varphi(x)) dU = 0. \quad (3.5)$$

As volume  $U$  is arbitrary, therefore, if  $\varphi(x)$  and  $\operatorname{div} W$  are continuous functions of point  $x = (x_1, x_2, x_3)$ , then (3.5) yields

$$W = \varphi(x). \quad (3.6)$$

Substituting here expression (3.1) for the vector of heat stream  $W$ , the following equation for the stationary temperature  $V = V(x)$  is obtained:

$$LV = \text{div}(K \text{grad } V) = -\varphi(x), \tag{3.7}$$

or in the equivalent form,

$$\sum_{\alpha=1}^3 (K(x)V(x)_{,x_\alpha})_{,x_\alpha} = -\varphi(x), \tag{3.8}$$

where  $K$  is the function of point  $x = (x_1, x_2, x_3)$ .

Heat transfer equation (3.8) is obtained under an assumption of isotropy of the heat transfer process. If a heat coefficient depends on direction and is a tensor (an isotropic medium), then, instead of (3.8), the following equation is taken:

$$\sum_{\alpha_1\beta=1}^3 (K_{\alpha\beta}V(x)_{,x_\beta})_{,x_\alpha} = -\varphi(x). \tag{3.9}$$

If  $K_{\alpha\beta} \equiv 0$  and  $\alpha \neq \beta$ , then (3.9) reads

$$\sum_{\alpha=1}^3 (K_\alpha V(x)_{,x_\alpha})_{,x_\alpha} = -\varphi(x). \tag{3.10}$$

Equation (3.8) holds for all internal points of space  $G$ . Additional conditions of the following form are attached to boundary  $\Gamma$ :

- (1) temperature is given:  $V(x) = g(x)$  for  $x \in \Gamma$ ;
- (2) heat stream is given:  $K(\partial V/\partial n) = g(x)$  for  $x \in \Gamma$ ;
- (3) heat transfer of Newton's rule is applied:

$$K \frac{dV}{dn} = \varkappa V + g(x), \quad x \in \Gamma, \varkappa = \varkappa(x) > 0. \tag{3.11}$$

Consider the boundary value problem for the equation of the stationary heat transfer in a rectangular parallelepiped of a three-dimensional space:

$$\bar{G} = \{0 \leq x_\alpha \leq l_\alpha, \alpha = 1, 2, 3\}. \tag{3.12}$$

In this case, the method of finite differences is used; the construction of the difference equation on the mesh with a variable step is realized through the method of functional approximation [5]. Consider the boundary value conditions (1), (2), and (3) in various states, and, if necessary, symmetry  $q$  of solution is accounted:

$$\sum_{\alpha=1}^3 (K_\alpha(X)V(x)_{,x_\alpha})_{,x_\alpha} = -\varphi(x), \quad x \in \bar{G}; \tag{3.13}$$

$$\begin{aligned} K_\alpha(x)V(x)_{,x_\alpha} &= \varkappa_{-\alpha}(x)(V(x) - V_0(x)), & x_\alpha &= 0, \\ -K_\alpha(x)V(x)_{,x_\alpha} &= \varkappa_{+\alpha}(x)(V(x) - V_0(x)), & x_\alpha &= l_\alpha, \end{aligned} \tag{3.14}$$

where  $K_\alpha(x)$  is the heat transfer coefficient in direction  $x_\alpha$ ,  $x_{\pm\alpha}(x)$  is the heat exchange coefficient with surrounding medium on walls  $x_\alpha = 0, l_\alpha$ , and  $V_\alpha(x)$  is the temperature of surrounding medium.

It is known that a solution to the boundary value problem (3.13), (3.14) satisfies the functional minimum:

$$\begin{aligned}
 I(u) = & \iiint_G \left( \sum_{\alpha=1}^3 k_\alpha V_{,x_\alpha}^2 - 2V\varphi \right) dx \\
 & + \sum_{\alpha=1}^3 \int_0^{l_\beta} \int_0^{l_\gamma} \left( (V^2 - 2g_{-\alpha}V) |_{x_\alpha=0} + (V^2 - 2g_{+\alpha}V) |_{x_\alpha=l_\alpha} \right) dx_\beta dx_\gamma, \quad \alpha \neq \beta \neq \gamma,
 \end{aligned}
 \tag{3.15}$$

where  $g_{\pm\alpha}(x) = \alpha_{\pm\alpha}(x)V_0(x)$ .

Indeed, varying functional (3.15) with respect to  $V$ ,  $V_{,x_\alpha}$ , integrating by parts the terms with  $\delta V_{,x_\alpha}$ , and comparing variations of the functional to zero, relations (3.4) and (3.14) are obtained as the necessary conditions to minimize functional (3.15).

Further, instead of  $V$  and  $K$ ,  $v$  and  $k$  will be used.

We introduce the following fundamental mesh:

$$\bar{\omega} = \prod_{\alpha=1}^3 \bar{\omega}_\alpha,
 \tag{3.16}$$

where  $\bar{\omega}_\alpha = \{x_{\alpha i}, i = 0, 1, 2, \dots, N_\alpha - 1, N_\alpha\}$ —mesh on interval  $[0, l_\alpha]$ . The set of interval mesh nodes is denoted by

$$\omega = \prod_{\alpha=1}^3 \omega_\alpha,
 \tag{3.17}$$

where  $\omega_\alpha = \{x_{\alpha i}, i = 1, 2, \dots, N_\alpha - 1\}$ .

We introduce the following notations:

$$\begin{aligned}
 h_{\alpha i} &= x_{\alpha i} - x_{\alpha i-1}, \quad i = 1, 2, \dots, N_\alpha, \\
 \tilde{h}_{\alpha i} &= \begin{cases} 0, & 5_{\alpha 1}, i = 0, \\ 0, & 5h_{\alpha i-1} + h_{\alpha i}, i = 1, 2, \dots, N_\alpha - 1, \\ 0, & 5h_{\alpha N_\alpha}, i = N_\alpha, \end{cases} \\
 v_{i,j,k} &= v(x_{1i}, x_{2j}, x_{3k}), \\
 k_1^{i\pm 0,5,j,k} &= 0,5(k_1^{i,j,k} + k_1^{i\pm 1,j,k}), \\
 k_2^{i,j,\pm 0,5,k} &= 0,5(k_2^{i,j,k} + k_2^{i,j,\pm 1,k}), \\
 k_3^{i,j,k\pm 0,5} &= 0,5(k_3^{i,j,k} + k_3^{i,j,k\pm 1}).
 \end{aligned}
 \tag{3.18}$$



All integrals occurring in functional (3.15) are substituted by quadratures. If the under-integral function includes a partial derivative, then the center triangle formula is applied with respect to the corresponding variable (in other cases, the trapezoid rule is used).

Derivatives are approximated through difference approximations, for instance, by  $v_{,x1}(x_{1i} - 0, 5h_{1i}x_{2j}, x_{3k}) = (v_{i,j,k} - v_{i-1,j,k})/h_{1i}$ .

In what follows, the function with  $n$  variables  $v_{i,j,k}$  is obtained:

$$\begin{aligned}
 I_n(v) = & \sum_{i=1}^{N_1} \sum_{j=0}^{N_2} \sum_{k=0}^{N_3} \frac{k_1^{i-0,5,j,k} (v_{i,j,k} - v_{i-1,j,k})^2 \tilde{h}_{2j} \tilde{h}_{3k}}{h_{1i}} \\
 & + \sum_{i=0}^{N_1} \sum_{j=1}^{N_2} \sum_{k=0}^{N_3} \frac{k_2^{i,j-0,5,k} (v_{i,j,k} - v_{i,j-1,k})^2 \tilde{h}_{1i} \tilde{h}_{3k}}{h_{2j}} \\
 & + \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} \sum_{k=1}^{N_3} \frac{k_3^{i,j,k-0,5} (v_{i,j,k} - v_{i,j,k-1})^2 \tilde{h}_{1i} h_{2i}}{h_{3k}} \\
 & + \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} \sum_{k=0}^{N_3} \varphi_{i,j,k} v_{i,j,k} \tilde{h}_{1i} \tilde{h}_{2j} \tilde{h}_{3k} \\
 & + \sum_{\alpha=1}^3 \sum_{j=0}^{N_\beta} \sum_{k=0}^{N_\gamma} \alpha_{-\alpha} v(x_{\alpha 0}, x_{\beta j}, x_{\gamma k})^2 - 2g_{-\alpha} v(x_{\alpha 0}, x_{\beta j}, x_{\gamma k}) \\
 & \quad + \alpha_{+\alpha} v(x_{\alpha N_\alpha}, x_{\beta j}, x_{\gamma k})^2 - 2g_{+\alpha} v(x_{\alpha N_\alpha}, x_{\beta j}, x_{\gamma k}) \tilde{h}_{\beta j} \tilde{h}_{\gamma k}, \quad \alpha \neq \beta \neq \gamma, \\
 n = & \prod_{\alpha=1}^3 (N_\alpha + 1).
 \end{aligned} \tag{3.19}$$

As known, the minimum of this function is achieved at the point where its partial derivatives are equal to zero with respect to  $v_{i,j,k}$ .

We construct a system of difference equations of order  $n$  to define values  $v_{i,j,k}$  realizing the minimum of the functional (3.15) and being a solution to the boundary value problem (3.13), (3.14). The following four cases are considered.

(1) Internal mesh nodes of space  $\bar{G}$ .

Terms (3.19) including  $v_{i,j,k}$  are reported below:

$$\begin{aligned}
 S_1 = & \left( \frac{k_1^{i-0,5,j,k} (v_{i,j,k} - v_{i-1,j,k})^2}{h_{1i}} + \frac{k_1^{i+0,5,j,k} (v_{i+1,j,k} - v_{i,j,k})^2}{h_{1i+1}} \right) \tilde{h}_{2j} \tilde{h}_{3k} \\
 & + \left( \frac{k_2^{i,j-0,5,k} (v_{i,j,k} - v_{i,j-1,k})^2}{h_{2j}} + \left( \frac{k_2^{i,j+0,5,k} (v_{i,j+1,k} - v_{i,j,k})^2}{h_{2j+1}} \right) \tilde{h}_{1i} \tilde{h}_{3k} \right. \\
 & \quad \left. + \left( \frac{k_3^{i,j,k-0,5} (v_{i,j,k} - v_{i,j,k-1})^2}{h_{3k}} \right) \frac{k_3^{i,j,k+0,5} (v_{i,j,k+1} - v_{i,j,k})^2}{h_{3k+1}} \right) \tilde{h}_{1i} \tilde{h}_{2j} \\
 & - 2\varphi_{i,j,k} v_{i,j,k} \tilde{h}_{1i} \tilde{h}_{2j} \tilde{h}_{3k}.
 \end{aligned} \tag{3.20}$$

By differentiating  $S_1$  with respect to  $v_{(i,j,k)}$  and comparing its derivative to zero, the following difference equation is obtained at node  $\{i, j, k\}$  of the mesh owing to the division by  $2h_1h_2h_3k$ :

$$\begin{aligned} & \frac{(k_1^{i+0,5,j,k}(v_{i+1,j,k} - v_{i,j,k})/h_{1i+1} - k_1^{i-0,5,j,k}(v_{i,j,k} - v_{i-1,j,k})/h_{1i})}{\tilde{h}_{1i}} \\ & + \frac{(k_2^{i,j+0,5,k}(v_{i,j+1,k} - v_{i,j,k})/h_{2j+1} - k_2^{i,j-0,5,k}(v_{i,j,k} - v_{i,j-1,k})/h_{2j})}{\tilde{h}_{2j}} \\ & + \frac{(k_3^{i,j,k+0,5}(v_{i,j,k+1} - v_{i,j,k})/h_{3k+1} - k_3^{i,j,k-0,5}(v_{i,j,k} - v_{i,j,k-1})/h_{3k})}{\tilde{h}_{3k}} \\ & = -\varphi_{i,j,k}, \quad 1 \leq i \leq N_1 - 1, \quad 1 \leq j \leq N_2 - 1, \quad 1 \leq k \leq N_3. \end{aligned} \tag{3.21}$$

(2) Node of the parallelepiped wall.

$$i = 0, \quad 1 \leq j \leq N_2 - 1, \quad 1 \leq k \leq N_3 - 1. \tag{3.22}$$

Terms (3.19) (including variable  $v_{i,j,k}$  for  $i = 0$ ) follow:

$$\begin{aligned} S_2 &= \frac{k_1^{i+0,5,j,k}(v_{i+1,j,k} - v_{i,j,k})^2 \tilde{h}_{2j} \tilde{h}_{3k}}{h_{1i+1}} \\ &+ \left( \frac{k_2^{i,j-0,5,k}(v_{i,j,k} - v_{i,j-1,k})^2}{h_{2j}} + \frac{k_2^{i,j+0,5,k}(v_{i,j+1,k} - v_{i,j,k})^2}{h_{2j+1}} \right) \tilde{h}_{1i} \tilde{h}_{3k} \\ &+ \left( \frac{k_3^{i,j,k-0,5}(v_{i,j,k} - v_{i,j,k-1})^2}{h_{3k}} + \frac{k_3^{i,j,k+0,5}(v_{i,j,k+1} - v_{i,j,k})^2}{h_{3k+1}} \right) \tilde{h}_{1i} \tilde{h}_{2j} \\ &- 2\varphi_{i,j,k} v_{i,j,k} \tilde{h}_{1i} \tilde{h}_{2j} \tilde{h}_{3k} + (\alpha_{-1} v_{i,j,k}^2 - 2g_{-1} v_{i,j,k}) \tilde{h}_{2j} \tilde{h}_{3k}. \end{aligned} \tag{3.23}$$

By differentiating  $S_2$  with respect to  $v_{i,j,k}$  and comparing the corresponding derivative to zero, the following difference equation (after division by  $2\tilde{h}_{1i}\tilde{h}_{2j}\tilde{h}_{3k}$ ) is obtained at the parallelepiped node  $x_1 = 0$ :

$$\begin{aligned} & \frac{(k_1^{i+0,5,j,k}(v_{i+1,j,k} - v_{i,j,k})/h_{1i+1} - \alpha_{-1} v_{i,j,k})}{\tilde{h}_{1i}} \\ & + \frac{(k_2^{i,j+0,5,k}(v_{i,j+1,k} - v_{i,j,k})/h_{2j+1} - k_2^{i,j-0,5,k}(v_{i,j,k} - v_{i,j-1,k})/h_{2j})}{\tilde{h}_{2j}} \\ & + \frac{(k_3^{i,j,k+0,5}(v_{i,j,k+1} - v_{i,j,k})/h_{3k+1} - k_3^{i,j,k-0,5}(v_{i,j,k} - v_{i,j,k-1})/h_{3k})}{\tilde{h}_{3k}} \\ & = -\varphi_{i,j,k} - \frac{g_{-1}}{\tilde{h}_{1i}}. \end{aligned} \tag{3.24}$$

Difference equations at other parallelepiped wall nodes are constructed analogously.

(3) Mesh nodes belong to the parallelepiped

$$i = j = 0, \quad 1 \leq k \leq N_3 - 1. \tag{3.25}$$

By proceeding in a way similar to that applied in the previous case, the following difference equation is obtained:

$$\begin{aligned} & \left( \frac{k_1^{i+0,5,j,k} (v_{i+1,j,k} - v_{i,j,k})}{h_{1i+1}} - \alpha_{-1} v_{i,j,k} \right) \tilde{h}_{1i} \\ & + \frac{(k_2^{i,j+0,5,k} (v_{i,j+1,k} - v_{i,j,k})/h_{2j+1} - \alpha_{-2} v_{i,j,k})}{\tilde{h}_{2j}} \\ & + \frac{(k_3^{i,j,k+0,5} (v_{i,j,k+1} - v_{i,j,k})/h_{3k+1} - k_3^{i,j,k-0,5} (v_{i,j,k} - v_{i,j,k-1})/h_{3k})}{\tilde{h}_{3k}} \\ & = -\varphi_{i,j,k} - \frac{g-1}{\tilde{h}_{1i}} - \frac{g-2}{\tilde{h}_{2j}}. \end{aligned} \tag{3.26}$$

(4) A node of the mesh coincides with the parallelepiped vertex

$$i = 0, N_1, \quad j = 0, N_2, \quad k = 0, N_3. \tag{3.27}$$

Proceeding again as in the previous case, one arrives at the following corresponding equations: for  $i = j = k = 0$ ,

$$\begin{aligned} & \frac{(k_1^{i+0,5,j,k} (v_{i+1,j,k} - v_{i,j,k})/h_{1i+1} - \alpha_{-1} v_{i,j,k})}{\tilde{h}_{1i}} \\ & + \frac{(k_2^{i,j+0,5,k} (v_{i,j+1,k} - v_{i,j,k})/h_{2j+1} - \alpha_{-2} v_{i,j,k})}{\tilde{h}_{2j}} \\ & + \frac{(k_3^{i,j,k+0,5} (v_{i,j,k+1} - v_{i,j,k})/h_{3k+1} - \alpha_{-3} v_{i,j,k})}{\tilde{h}_{3k}} \\ & = -\varphi_{i,j,k} - \frac{g-1}{\tilde{h}_{1i}} - \frac{g-2}{\tilde{h}_{2j}} - \frac{g-3}{\tilde{h}_{3k}}; \end{aligned} \tag{3.28}$$

for  $i = N_1, j = N_2$ , and  $k = N_3$ ,

$$\begin{aligned} & \frac{(-\alpha_{-1} v_{i,j,k} - k_1^{i-0,5,j,k} (v_{i,j,k} - v_{i-1,j,k})/h_{1i})}{\tilde{h}_{1i}} \\ & + \frac{(-\alpha_{-2} v_{i,j,k} - k_2^{i,j-0,5,k} (v_{i,j,k} - v_{i,j-1,k})/h_{2j})}{\tilde{h}_{2j}} \\ & + \frac{(-\alpha_{-3} v_{i,j,k} - k_3^{i,j,k-0,5} (v_{i,j,k} - v_{i,j,k-1})/h_{3k})}{\tilde{h}_{3k}} \\ & = -\varphi_{i,j,k} - \frac{g+1}{\tilde{h}_{1i}} - \frac{g+2}{\tilde{h}_{2j}} - \frac{g+3}{\tilde{h}_{3k}}. \end{aligned} \tag{3.29}$$

Thus, all possible cases occurring while formulating the difference equations on the mesh with variable step for the third boundary value problem have been discussed.

We discuss briefly the peculiarities involved in the formulation of difference equations in the case when the following Dirichlet conditions are given on some parallelepiped walls:

$$v(x)|_{x_\alpha=0, l_\alpha} = v_u(x), \quad (3.30)$$

where  $v_u(x)$  is the given function.

One of the possible ways to include a condition of such a type is to formulate the following corresponding equation associated with nodes:

$$v_{i,j,k} = v_u. \quad (3.31)$$

The latter one allows for construction of both an algorithm and computer programs universal with respect to the boundary condition type. Indeed, the drawback of the approach is associated with conservation of the order of the difference equation system.

Furthermore, if it is known a priori that the solution to problem (3.13), (3.14) possesses one, two, or three plane symmetries  $x_\alpha = l_\alpha/2$ , then these conditions are recommended to be included while difference equations are being formulated, since they essentially decrease the order of the investigated system.

If the solution is symmetric with respect to, say, the plane  $x_1 = \ell_1/2$ , then we take

$$\tilde{h}_{1N_1} = h_{1N_1}, \quad k_1^{N_1+0,5,j,k} = k_1^{N_1-0,5,j,k}. \quad (3.32)$$

In what follows, in the difference equations formulated for the nodes of the plane  $x_\alpha = \ell_\alpha/2$ , accounting for the symmetry property

$$v\left(\frac{\ell_1}{2} + h_{1N_1}, x_2, x_3\right) = v\left(\frac{\ell_1}{2} - h_{1N_1}, x_2, x_3\right), \quad (3.33)$$

the coefficient standing by the unknown  $v_{i-1,j,k}$  is doubled. The wall node is treated as an internal one in this case.

The solution to the boundary value problem (3.13), (3.14) consists of two steps: formulation of the system of difference equations with a band-type compressed matrix having a regular structure (in the operating storage, only nonlinear matrix diagonals are conserved), and solution to the obtained system using the elimination method for equations [3].

We discuss the computations of the first stage. In order to identify a particular problem, the following items are required: node numbers  $N_\alpha$  with respect to each direction  $x_\alpha$ ,  $\alpha = 1, 2, 3$ ; packets of steps of the mesh  $h_{1i}$ ,  $h_{2j}$ ,  $h_{3k}$ ; formulas for  $k_\alpha(x)$ ,  $g_{\pm\alpha}(x)$ ,  $\varphi(x)$  computations, and  $v_u(x)$  (if necessary); values of symmetry indicators  $IS$ ,  $JS$ ,  $KS$ ; packet of six elements to formulate the type of the boundary condition on each of the six parallelepiped walls; and values of  $\chi_{\pm\alpha}(x)$ .

Having this information, the operator coefficients  $k_1^{i,j,k}$ ,  $k_1^{i-0,5,j,k}$  are computed and the matrix of difference equations is formulated due to formulas (3.21)–(3.29). Analogous operations are carried out with  $k_2(x)$ ,  $k_3(x)$ .

In what follows, the matrix of coefficients of the difference equations, which corresponds to the third boundary value problem, is obtained. Then free terms of the system of equations are computed, and (if necessary) the equations involving the Dirichlet conditions are formulated.

Recall that when solving numerically the boundary value problems, the reality estimation of the computer-yielded results is in the focus of attention. Therefore, the correct derivation of appropriate relations together with suitable formulation of the variational problem should be emphasized, and then an application of the exact method of solution to the system of difference equations with double precision should be carried out.

During analysis, a series of computations to verify the mentioned reliability has been carried out.

(1) Model problem. The possibilities of both the developed algorithm and the system subroutines are checked through solution to problem (3.13), (3.14), with

$$\begin{aligned} k_1 &= 1 + \frac{2}{3c(\tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3)}, \\ k_2 &= 1 + \frac{2}{3c(3/4 - \tilde{x}_1 - \tilde{x}_2 - \tilde{x}_3)}, \\ k_3 &= 1 + \frac{2}{3c(1/2 - \tilde{x}_1 - \tilde{x}_2 + \tilde{x}_3)}, \end{aligned} \tag{3.34}$$

where

$$\tilde{x}_\alpha = \left(x_\alpha - \frac{1}{2}\right)^2, \quad \alpha = 1, 2, 3. \tag{3.35}$$

The exact solution reads

$$v_u(x) = \frac{(1/2 - \tilde{x}_1 - \tilde{x}_2)x_3^2}{2}, \tag{3.36}$$

and hence

$$\begin{aligned} \varphi(x) &= -x_3^2 \left( 2 + \frac{4}{3c(3/8 + \tilde{x}_1 - \tilde{x}_2)} \right) \\ &+ \left( \frac{1}{2} - \tilde{x}_1 - \tilde{x}_2 \right) \left( 1 + \frac{2}{3c(1/2 - \tilde{x}_1 - \tilde{x}_2 + \tilde{x}_3 + 2x_3(x_3 - 1/2))} \right), \\ g_{-2}(x_1, 0, x_3) &= -\frac{3}{4} - \tilde{x}_1 + \frac{2}{3c(1/2 - \tilde{x}_1 - x_2)x_3^2/2}, \\ g_{-3}(x_1, 0, x_3) &= 0, \\ g_{+3}(x_1, x_2, 1) &= \left( \frac{1}{2} - \tilde{x}_1 - \tilde{x}_2 \right) \left( \frac{3}{2} + \frac{3}{2c(3/4 - \tilde{x}_1 - \tilde{x}_2)} \right) \end{aligned} \tag{3.37}$$

assuming that  $\alpha_{\pm\alpha}(x) = 1$ .

We introduce the Dirichlet condition on the wall  $x_1 = 0$ . The symmetry of solution is taken into account with respect to the plane  $x_1 = 1/2, x_2 = 1/2$ . The solution is defined in the space  $\bar{G} = \{0 \leq x_\alpha \leq 1, \alpha = 1, 2, 3\}$ .

Table 3.1. Numerical solution of the model problems at point:  $x_1 = 0.5, x_2 = 0.5, x_3 = x_{3k}$ .

$k$	$x_{3k}$	Mesh			Exact solution
		$3 \times 3 \times 5$	$5 \times 5 \times 5$	$9 \times 9 \times 9$	
1	0	0	0	0	0
2	0.25	0.0112	0.0120	0.0147	0.0156
3	0.5	0.504	0.0523	0.0599	0.0625
4	0.75	0.120	0.122	0.136	0.141
5	1.0	0.223	0.0224	0.243	0.250

Table 3.2. Numerical solution of the model problems at point:  $x_{1k} = x_{2k} = x_{3k}$ .

$k$	$x_{3k}$	Mesh			Exact solution
		$3 \times 3 \times 3$	$5 \times 5 \times 5$	$9 \times 9 \times 5$	
1	0	0	0	0	0
2	0.25	—	0.0038	0.0060	0.0068
3	0.5	0.378	0.0375	0.0444	0.0469
4	0.75	—	0.114	0.127	0.132

In Table 3.1, the results of the solution to this problem for  $c = 2048$  are reported for the mesh with a constant step. An analysis of the given results yields the conclusion that during the change of the mesh step, the approximated solution converges to the exact one. Variation of the operator coefficients in the considered case is high, that is,

$$1 \leq k_\alpha(x) \leq 1024, \quad \alpha = 1, 2, 3. \tag{3.38}$$

Notice that for  $c = 0$ , the numerical solution with accuracy greater than five meaningful digits overlaps with the exact solution.

(2) Problem without a source. The convergence of the solution to the problem without a source (owing to the decreasing mesh step) is testified through the results included in Table 3.2. The following parameters are taken:

$$k_1 = k_2 = k_3 = 207, \quad \varphi(x) = 0, \tag{3.39}$$

and the following boundary conditions are applied:

$$\begin{aligned} v(0, x_2, x_3) = 0, \quad v_{,x_2}(x_1, 0, x_3) = 0, \\ k_3 v_{,x_3}(x_1, x_2, 0) = 10(v - 100), \quad k_3 v_{,x_3}(x_1, x_2, l_3) = 10v. \end{aligned} \tag{3.40}$$

The following symmetry solution with respect to the planes is taken into account:

$$x_1 = \frac{l_1}{2}, \quad x_2 = \frac{l_2}{2}; \quad l_1 = l_2 = 1m, \quad l_3 = 0.02m. \tag{3.41}$$

Table 3.3. Numerical solution to the problem without source.

$k$	$x_{1k}$	Mesh		Recipes
		$5 \times 5 \times 5$	$9 \times 9 \times 5$	
1	0	0	0	—
2	0.125	9.17	9.2	$x_{1k} = x_{2k}$
3	0.25	15.3	15.3	$x_3 = \ell_3$
4	0.375	18.8	18.8	—
5	0.5	19.9	20.0	—

Owing to the analysis of the results reported in Table 3.3, it is clear that without sources, the numerical solution is obtained with an efficient accuracy on the mesh  $5 \times 5$ .

(3) Problem with a point source. The convergence of the solution to the problem with a point source obtained during the decrease of the mesh step is considered on the basis of the following problem:

$$k_1 = k_2 = k_3 = 1, \quad \varphi(x) = \begin{cases} 4 \cdot 10^6, & \text{for } x_1 = x_2 = \frac{\ell_1}{2}, x_3 = \ell_3, \\ 0, & \text{otherwise,} \end{cases} \quad (3.42)$$

where the attached boundary conditions are as follows:

$$\begin{aligned} v(0, x_2, x_3) = 0, & \quad v_{,x_2}(x_1, 0, x_3) = 10v, \\ v_{,x_3}(x_1, x_2, 0) = 0, & \quad v_{,x_3}(x_1, x_2, \ell_3) = 0; \\ \ell_1 = \ell_2 = 1, & \quad \ell_3 = 0.02. \end{aligned} \quad (3.43)$$

The following symmetry condition with respect to the planes

$$x_1 = \frac{\ell_1}{2}, \quad x_2 = \frac{\ell_2}{2} \quad (3.44)$$

is applied.

In Tables 3.4 and 3.5, the results of the solution to this problem on four meshes,  $3 \times 3 \times 5$ ,  $5 \times 5 \times 5$ ,  $9 \times 9 \times 5$ , and  $17 \times 17 \times 5$ , are given. For the three partitions, the mesh step (in the vicinity of the node with the source) is taken to be constant and equal to the step of partition  $17 \times 17 \times 5$ . In all four cases, the constant volume power is achieved. The changeable step allows application of the mentioned actions in a relatively simple manner by introduction of additional nodes with respect to  $x_1$ ,  $x_2$  in the first three cases, that is, the solution is obtained for the shells  $4 \times 4 \times 5$ ,  $6 \times 6 \times 5$ , and  $10 \times 10 \times 5$ .

Owing to the given results, stable convergence is achieved during the mesh step variation. Maximal differences are observed at the node with the source, but the difference is less than 1% for two neighboring partitions.

Table 3.4. Numerical results of the problem with point source.

$k$	$k_{3k}$	Mesh				Recipes
		$3 \times 3 \times 5$	$5 \times 5 \times 5$	$9 \times 9 \times 5$	$17 \times 17 \times 5$	
1	0	264.5	300.9	321.0	324.3	—
2	0.25	269.5	305.7	325.9	329.2	$x_1 = 0.5$
3	0.5	284.4	320.6	340.9	344.2	$x_2 = 0.5$
4	0.75	310.9	347.2	367.5	370.8	—
5	1	351.7	388.0	408.3	411.6	—

Table 3.5. Numerical results of the problem with point source.

$k$	$k_{1k}$	Mesh				Recipes
		$3 \times 3 \times 5$	$5 \times 5 \times 5$	$9 \times 9 \times 5$	$17 \times 17 \times 5$	
1	0	0	0	0	0	—
2	0.25	—	12.05	12.25	12.31	$x_{2k} = x_{1k}$
3	0.5	36.6	38.9	39.6	39.8	$x_{3k} = \ell_3$
4	0.75	—	90.0	92.0	92.5	—

(4) Comparison of solutions obtained through various methods. Finally, we present the results of computation through the finite difference method for two problems for which both analytical and boundary value solutions are known.

In Table 3.6, the results of the solution to the first boundary value problem for the Laplace equation for the cube with ribs of unit length are reported.

The boundary conditions are as follows:

$$\begin{aligned}
 v(-0.5, x_2, x_3) &= 2, & v(0.5, x_2, x_3) &= 1, \\
 v(x_1, -0.5, x_3) &= 0, & v(x_1, 0.5, x_3) &= 0, \\
 v(x_1, x_2, -0.5) &= 0, & v(x_1, x_2, 0.5) &= 0.
 \end{aligned}
 \tag{3.45}$$

In Table 3.7, the results of the solution to the Laplace equation with a hybrid boundary condition for the rectangular parallelepiped are reported. The boundary conditions are

$$\begin{aligned}
 v(-0.5, x_2, x_3) &= 10, & (v_{,x_1} + 5v)|_{x_1=0.5} &= 0, \\
 (v_{,x_2} + 5v)|_{x_2=\pm 1} &= 0, & (v_{,x_3} + 5v)|_{x_3=\pm 1} &= 0.
 \end{aligned}
 \tag{3.46}$$

Notice that when numerical solutions are constructed, the symmetry with respect to two planes,  $x_2 = 0, x_3 = 0$ , is taken into account.

The reported results show that the solution obtained through the finite difference method during the mesh step decrease converges to the analytical solution, and its error is small for the appropriate partition.



Table 3.6. Solution to the first boundary value problem for Laplace equation.

$x_{1k}$	Boundary element method		Finite difference method			Analytical solution
	$N = 12$	$N = 24$	$5 \times 3 \times 3$	$9 \times 5 \times 5$	$17 \times 9 \times 9$	
-0.375	1.637	1.472	—	1.413	1.426	1.430
-0.250	1.044	0.979	0.926	0.953	0.963	0.967
-0.125	0.678	0.661	—	0.653	0.657	0.659
0	0.5	0.5	0.5	0.5	0.5	0.5
0.125	0.478	0.472	—	0.471	0.472	0.472
0.250	0.597	0.566	0.545	0.555	0.558	0.560
0.375	0.770	0.770	—	0.740	0.746	0.748

Table 3.7. Solution of the Laplace equation with hybrid boundary conditions ( $x_{3k} = x_{2k}$ ).

$x_{1k}$	$x_{2k}$	Boundary element method		Finite element method			Analytical solution
		$N = 24$	$N = 48$	$5 \times 3 \times 3$	$9 \times 5 \times 5$	$17 \times 9 \times 9$	
-0.25	0	7.387	7.282	7.155	7.234	7.257	7.259
0	0	4.827	4.84	4.724	4.804	4.829	4.837
0	0.5	3.745	3.843	3.821	3.835	3.841	3.843
0.25	0	2.816	2.843	2.778	2.825	2.839	2.844
0.25	0.25	2.612	2.658	—	2.645	2.655	2.658
0.25	0	2.0	2.073	2.105	2.094	2.091	2.089
0.25	0.75	1.050	1.144	—	1.202	1.186	1.180

The results of the numerical experiments yield the conclusion that our proposed method and the associated subroutines allow efficient numerical solution of three-dimensional boundary problems for the stationary heat transfer equation without any essential limitations. The same conclusion holds when the problems of plates and shells are considered.

#### 4. Computation of geometrically nonlinear nonhomogenous shallow shells with mixed boundary condition along their sides

One of the important computation problems connected with shallow nonhomogeneous and geometrically nonlinear shells concerns conditions of the boundary value problems variations along a supporting side.

In order to derive appropriate relations, a system consisting of two identical shells stiffened by ribs along two sides linked by a hinge is considered (see [Figure 4.1](#)).

The shell material is assumed to be isotropic but nonhomogeneous, that is, shear modulus  $G$  and Poisson coefficient  $\mu$  are functions of point coordinates  $x = (x_1, x_2, x_3)$ , or,

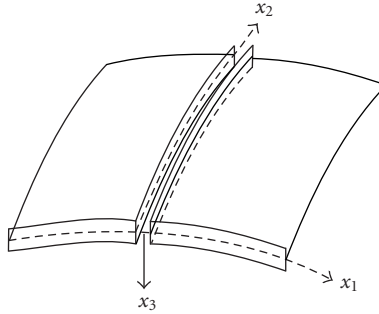


Figure 4.1. Investigated shell.

according to the theory of small elastic-plastic deformations, they depend on the stress-strain material state at the considered point.

We introduce the following notation:  $x_1, x_2, x_3$ —rectangular system of coordinates (Figure 4.1);  $\alpha, b, h$ —shell dimension and thickness, respectively;  $k_1, k_2$ —curvatures;  $u, v, w$ —components of displacements of mean surface point;  $q$ —transversal load intensity.

Following the Kirchhoff-Love hypothesis, the components of shell and rib deformation read

$$e_{ii} = \varepsilon_{ii} + \chi_{ii}x_3 \quad (i = 1, 2), \quad e_{12} = \varepsilon_{12} + \chi_{12}x_3, \quad (4.1)$$

where

$$\begin{aligned} \varepsilon_{11} &= u_{,x_1} - k_1 w + \frac{w_{,x_1}^2}{2}, \\ \varepsilon_{22} &= v_{,x_2} - k_2 w + \frac{w_{,x_2}^2}{2}, \quad \varepsilon_{12} = U_{,x_2} + v_{,x_1} + w_{,x_1} w_{,x_2}, \end{aligned} \quad (4.2)$$

$\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{12}$ —length and stream deformations of the mean surface;

$$\chi_{11} = -w_{x_1 x_1}, \quad \chi_{22} = -w_{x_2 x_2}, \quad \chi_{12} = -w_{x_1 x_2}, \quad (4.3)$$

where  $\chi_{11}, \chi_{22}, \chi_{12}$  are variations of curvature and torsion.

In accordance with the Hook principle for a flat-strain state and (4.1), the coupling between stressed  $\sigma_{ik}$  and deformation  $e_{ik}$  is presented in the following form:

$$\begin{aligned} \sigma_{11} &= \frac{2G}{1-\nu} [\varepsilon_{11} + \chi_{11}x_3 + \mu(\varepsilon_{22} + \chi_{22}x_3)] \quad (1 \leftrightarrow 2), \\ \sigma_{12} &= G(\varepsilon_{12} + 2\chi_{12}x_3), \end{aligned} \quad (4.4)$$

where, here and further,  $(x \leftrightarrow y)$  denotes the change of index  $x$  to  $y$ .

Integration of (4.4) with respect to  $x_3$  yields the following expressions for forces and movements:

$$T_{11} = c_{00}\varepsilon_{11} + c_{10}\varepsilon_{22} + c_{01}\chi_{11} + c_{11}\chi_{22}, \tag{4.5}$$

$$T_{22} = c_{00}\varepsilon_{11} + c_{00}\varepsilon_{22} + c_{11}\chi_{11} + c_{01}\chi_{22}, \tag{4.6}$$

$$T_{12} = 0.5(c_{00} - c_{10})\varepsilon_{12} + (c_{01} - c_{11})\chi_{12}, \tag{4.7}$$

$$M_{11} = c_{01}\varepsilon_{11} + c_{11}\varepsilon_{22} + c_{02}\chi_{11} + c_{12}\chi_{22}, \tag{4.8}$$

$$M_{22} = c_{11}\varepsilon_{11} + c_{01}\varepsilon_{22} + c_{12}\chi_{11} + c_{02}\chi_{22}, \tag{4.9}$$

$$M_{12} = 0.5(c_{01} - c_{11})\varepsilon_{12} + (c_{02} - c_{12})\chi_{12}, \tag{4.10}$$

where

$$c_{ik}(x_1, x_2) = \int_{-h/2}^{h/2} \frac{2G(x)\mu^i(x)x_3^k}{1 - \mu(x)} dx^3 \quad (i = 0, 1; k = 0, 1, 2). \tag{4.11}$$

The deformation energy of the considered system reads

$$\begin{aligned} \Xi = 0.5 \left[ \int_{-\alpha}^0 \int_0^b A dx_1 dx_2 + \int_0^\alpha \int_0^b A dx_1 dx_2 + \int_{-\alpha}^0 (H_1^- + P_1^-) \Big|_{x_2=0} dx_1 \right. \\ \left. + \int_0^\alpha (H_1^+ + P_1^+) \Big|_{x_2=0} dx_1 + \int_0^b (H_2^- + H_2^+ + P) \Big|_{x_1=0} dx_2 \right], \end{aligned} \tag{4.12}$$

where

$$\begin{aligned} A &= T_{11}\varepsilon_{11} + T_{22}\varepsilon_{22} + T_{12}\varepsilon_{12} + M_{11}\chi_{11} + M_{22}\chi_{22} + 2M_{12}\chi_{12}, \\ H_1^\pm &= D_1^\pm \varepsilon_{11}^2 + A_1^\pm w_{,x_1 x_1}^2 + B_1^\pm v_{,x_1 x_1}^2 + C_1^\pm w_{,x_1 x_2}^2, \\ H_2^\pm &= D_2^\pm \varepsilon_{22}^2 + A_2^\pm w_{,x_2 x_2}^2 + B_2^\pm u_{,x_2 x_2}^2 + C_2^\pm w_{,x_1 x_2}^2, \\ P_1^\pm &= \alpha_1^\pm w_{,x_2}^2 + \beta_1^\pm w^2, \quad P = \alpha[w_{,x_1}]^2 + \beta w^2; \end{aligned} \tag{4.13}$$

$D_1^\pm, A_1^\pm, B_1^\pm, C_1^\pm$ —rib stiffness on elongation (compressing), bending in a vertical plane, bending in a horizontal plane, and torsion, respectively;  $\alpha_1^\pm, \alpha$ —stiffness of hinge rib-shell, rib-rib interaction, respectively,  $\beta_1^\pm, \beta$ —elasticity of support under the rib;  $[w_{,x_1}]$ —jump of derivative  $w_{,x_1}$  on the line  $x_1 = 0$ .

We introduce the following functional:

$$I(u, v, w) = \Xi - \int_{-\alpha}^0 \int_0^b qw dx_1 dx_2 - \int_0^\alpha \int_0^b qw dx_1 dx_2. \tag{4.14}$$

Varying (4.14) with respect to  $u, v, w$  and carrying out transformations with the help of integration by parts, one gets

$$\begin{aligned}
 \delta I = & \int_{-\alpha}^0 \int_0^b \bar{A} dx_1 dx_2 + \int_0^\alpha \int_0^b \bar{A} dx_1 dx_2 \\
 & + \int_{-\alpha}^0 (R_1|_{x_2=0,b} + (S_1^- + \beta_1^- w \delta w + \alpha_1^- w_{,x_2} \delta w_{,x_2})|_{x_2=0}) dx_1 \\
 & + \int_0^\alpha (R_1|_{x_2=0,b} + (S_1^+ + \beta_1^+ w \delta w + \alpha_1^+ w_{,x_2} \delta w_{,x_2})|_{x_2=0}) dx_1 \\
 & + \int_0^b (R_2|_{x_1=-\alpha,0} + R_2|_{x_1=0,\alpha} + (S_2^- + S_2^+ + \alpha[w_{,x_1}] \delta[w_{,x_1}] + \beta w \delta w)|_{x_1=0}) dx_2 \\
 & + Q_1^- \Big|_{\substack{x_1=-\alpha,0 \\ x_2=0}} + (Q_2^+ + Q_2^-) \Big|_{\substack{x_1=0 \\ x_2=0,b}} + Q_1^+ \Big|_{\substack{x_1=0,\alpha \\ x_2=b}} \\
 & - 2M_{12} \delta w \Big|_{\substack{x_1=-\alpha,0 \\ x_2=0,b}} - 2M_{12} \delta w \Big|_{\substack{x_1=0,\alpha \\ x_2=0,b}} = 0,
 \end{aligned} \tag{4.15}$$

where

$$\begin{aligned}
 \bar{A} = & -(T_{11,x_1} + T_{12,x_2}) \delta u - (T_{12,x_1} + T_{22,x_2}) \delta v \\
 & - (M_{11,x_1x_1} + M_{22,x_2x_2} + 2M_{12,x_1x_2} + k_1 T_{11} + k_2 T_{22} \\
 & + (T_{11} w_{,x_1} + T_{12} w_{,x_2})_{,x_1} + (T_{12} w_{,x_1} + T_{22} w_{,x_2})_{,x_2} - q) \delta w, \\
 R_1 = & T_{12} \delta u + T_{22} \delta v + (M_{22,x_2} + 2M_{12,x_1} + T_{22} w_{,x_2}) \delta w - M_{22} \delta w_{,x_2}, \\
 S_1^\pm = & -(D_1^\pm \varepsilon_{11})_{,x_1} \delta u + (B_1^\pm v_{,x_1x_1})_{,x_1x_1} \delta v + (-k_1 D_1^\pm \varepsilon_{11}) - (D_1^\pm \varepsilon_{11} w_{,x_1}) \\
 & + (A_1^\pm w_{,x_1x_1}) \delta w - (C_1^\pm w_{,x_1x_2}) \delta w_{,x_2}, \\
 Q_1^\pm = & D_1^\pm \varepsilon_{11} \delta - (B_1^\pm v_{,x_1x_1})_{,x_1} \delta v + (B_1^\pm v_{,x_1x_1}) \delta v_{,x_1} \\
 & + (D_1^\pm \varepsilon_{11} w_{,x_1} - (A_1^\pm w_{,x_1x_1})_{,x_1}) \delta w + A_1^\pm w_{,x_1x_1} \delta w_{,x_1} + C_1^\pm w_{,x_1x_2} \delta w_{,x_2} \\
 & (1 \longleftrightarrow 2, u \longleftrightarrow v, x_1 \longleftrightarrow x_2).
 \end{aligned} \tag{4.16}$$

We consider some of the relations that follow from the previous considerations.

Differential equations governing the equilibrium state of a nonhomogeneous shell with respect to displacements are obtained by comparison to zero of the coefficients standing by variations of  $u, v, w$  in doubled integral (4.15), and by taking into account relations (4.2), (4.6), and (4.9).

However, they are not explicitly given owing to their complexity. Boundary conditions for the side  $x_2 = 0$  follow from the examinations of the under-integral expressions. If the displacements on the contour are not given, then the boundary conditions are defined by comparison to zero of the coefficients standing by variations of  $u, v, w, w_{,x_2}$ .

Conditions for the coupling of solutions on the line  $x_1 = 0$  are derived from the following expression:

$$\begin{aligned}
& R_2|_{x_1=-0} - R_2|_{x_1=0} + (S_2^- + S_2^+ + \alpha[w_{,x_1}])\delta[w_{,x_1}] + \beta w \delta w|_{x_1=0} \\
& = (-[T_{11}] + (B_2 u_{,x_2 x_2})_{,x_2 x_2})\delta u + (-[T_{12}] - D_2 \varepsilon_{22})\delta v \\
& \quad + (-[T_{13}] - D_2 k_2 \varepsilon_{22} - (D_2 \varepsilon_{22} w_{,x_2})_{,x_2} + (A_2 w_{,x_2 x_2})_{,x_2 x_2} + \beta w)\delta w \\
& \quad + (-M_{11}^- - (C_2^- w_{,x_1 x_2}^-)_{,x_2} - \alpha[w_{,x_1}])\delta w_{,x_1}^- \\
& \quad + (M_{11}^+ - (C_2^+ w_{,x_1 x_2}^+)_{,x_2} + \alpha[w_{,x_1}])\delta w_{,x_1}^+ = 0,
\end{aligned} \tag{4.17}$$

where

$$\begin{aligned}
[T_{1i}] &= T_{1i}|_{x_1=0} - T_{1i}|_{x_1=-0} \quad (i = 1, 2, 3), \\
T_{13} &= M_{11,x_1} + 2M_{12,x_2} + T_{11}w_{,x_1}, \quad B_2 = B_2^- + B_2^+, \\
D_2 &= D_2^- + D_2^+, \quad A_2 = A_2^- + A_2^+.
\end{aligned} \tag{4.18}$$

This means that the following conditions should be satisfied on the line of the solution coupling:

$$\begin{aligned}
[T_{11}] &= (B_2 u_{,x_2 x_2})_{,x_2 x_2}, \quad [T_{12}] = -D_2 \varepsilon_{22}, \\
[T_{13}] &= -D_2 k_2 \varepsilon_{22} - (D_2 \varepsilon_{22} w_{,x_2})_{,x_2} + (A_2 w_{,x_2 x_2})_{,x_2 x_2} + \beta w, \\
[M_{11}] &= (C_2^+ w_{,x_1 x_2})_{,x_2} + (C_2^- w_{,x_1 x_2}^-)_{,x_2}, \\
\alpha[w_{,x_1}] &= -M_{11}^+ - M_{11}^- + (C_2^+ w_{,x_1 x_2})_{,x_2} + (C_2^- w_{,x_1 x_2}^-)_{,x_2},
\end{aligned} \tag{4.19}$$

as well as the conditions on the continuity of displacements  $[u] = [v] = [w] = 0$ . Taking  $k_2 = \varepsilon_{22} = u = v = T_{11}w_{,x_1} = 0$  in (4.19), the conditions for the solution coupling for a plate are obtained [5].

We determine agreement conditions in the shell angle  $x_1 = 0, x_2 = 0$ . For this purpose, the forms related to an angle are considered in (4.15), that is,

$$\begin{aligned}
& (-Q_1^+ - Q_2^+ - 2M_{12}\delta w)|_{x_1=0, x_2=0} \\
& = (-D_1^+ \varepsilon_{11} + (B_2^+ u_{,x_2 x_2})_{,x_2})\delta u + ((B_1^+ v_{,x_1 x_1})_{,x_1} - D_2^+ \varepsilon_{22})\delta v \\
& \quad + (-2M_{12} - D_1^+ \varepsilon_{11} w_{,x_1} + (A_1^+ w_{,x_1 x_1})_{,x_1} - D_2^+ \varepsilon_{22} w_{,x_2} + (A_2^+ w_{,x_2 x_2})_{,x_2})\delta w \\
& \quad - B_1^+ v_{,x_1 x_1} \delta v_{,x_1} - B_2^+ u_{,x_2 x_2} \delta u_{,x_2} - (A_1^+ w_{,x_1 x_1} C_2^+ w_{,x_1 x_2})\delta w_{,x_1} \\
& \quad - (A_2^+ w_{,x_2 x_2} + C_1^+ w_{,x_1 x_2})\delta w_{,x_2} = 0.
\end{aligned} \tag{4.20}$$

One may derive now, for instance, the following compatibility conditions in the angle:

$$\begin{aligned}
& -D_1^+ \varepsilon_{11} + (B_2^+ u_{,x_2 x_2})_{,x_2} = 0, \quad -D_2^+ \varepsilon_{22} + (B_1^+ v_{,x_1 x_1})_{,x_1} = 0, \\
& -2M_{12}^- + (A_1^+ w_{,x_1 x_1})_{,x_1} + (A_2^+ w_{,x_2 x_2})_{,x_2} - D_1^+ \varepsilon_{11} w_{,x_1} - D_2^+ \varepsilon_{22} w_{,x_2} = 0; \\
& \quad B_1^+ v_{,x_1 x_1} = 0, \quad B_2^+ u_{,x_2 x_2} = 0, \\
& \quad A_1^+ w_{,x_1 x_1} + C_2^+ w_{,x_1 x_2} = 0, \quad A_2^+ w_{,x_2 x_2} + C_1^+ w_{,x_1 x_2} = 0.
\end{aligned} \tag{4.21}$$

Notice that if the ribs are excluded, then only one compatibility condition  $M_{12} = 0$  remains.

We derive compatibility conditions for the point  $x_1 = 0$ , where the change of boundary conditions occurs. They are obtained from the terms of (4.15) with respect to the point  $x_1 = 0$ ,  $x_2 = 0$ , and the coefficients with index 2 are equal to zero (homogeneous shell):

$$\begin{aligned}
 (Q_1^- - Q_1^+) \Big|_{\substack{x=0 \\ x_2=0}} &= (D_1^- \varepsilon_{11}^- - D_1^+ \varepsilon_{11}^+) \delta u + ((B_1^+ v_{,x_1 x_1}^+),_{x_1} - (B_1^- v_{,x_1 x_1}^-),_{x_1}) \delta v \\
 &+ (-D_1^+ \varepsilon_{11}^+ w_{,x_1}^+ + D_1^- \varepsilon_{11}^- w_{,x_1}^- + (A_1^+ w_{,x_1 x_1}^+),_{x_1} - (A_1^- w_{,x_1 x_1}^-),_{x_1}) \delta w \\
 &+ B_1^- v_{,x_1 x_1}^- \delta v_{,x_1}^- - B_1^+ v_{,x_1 x_1}^+ \delta v_{,x_1}^+ + (C_1^- w_{,x_1 x_2}^- - C_1^+ w_{,x_1 x_2}^+) \delta w_{,x_2} \\
 &+ A_1^- w_{,x_1 x_1}^- \delta w_{,x_1}^- - A_1^+ w_{,x_1 x_1}^+ \delta w_{,x_1}^+ = 0.
 \end{aligned} \tag{4.22}$$

As an example, we consider the compatibility condition for a hinged support of the shell side  $x_2 = 0$  with different stiffness:

$$\begin{aligned}
 D_1^- \varepsilon_{11}^- - D_1^+ \varepsilon_{11}^+ &= 0, & (B_1^+ v_{,x_1 x_1}^+),_{x_1} - (B_1^- v_{,x_1 x_1}^-),_{x_1} &= 0, \\
 B_1^+ v_{,x_1 x_1}^+ &= 0, & C_1^- w_{,x_1 x_2}^- - C_1^+ w_{,x_1 x_2}^+ &= 0.
 \end{aligned} \tag{4.23}$$

Owing to relations (4.22), (4.23), the necessity of application of the compatibility condition appears only for the shells supported by ribs.

We introduce the following force function:

$$F_{,x_2 x_2} = T_{11}, \quad F_{,z_1 x_2} = -T_{12}, \quad F_{,x_1 x_1} = T_{22}, \tag{4.24}$$

and we derive the equations in a hybrid form with respect to force  $F$  and deflection  $w$ .

For this purpose, the deformation continuity equation and the third equation of (4.15) are applied to yield

$$\begin{aligned}
 \varepsilon_{11,x_2 x_2} + \varepsilon_{22,x_1 x_1} - \varepsilon_{12,x_1 x_2} + \nabla_k^2 w + 0,5L(w, w) &= 0, \\
 M_{11,x_1 x_1} + M_{22,x_2 x_2} + 2M_{12,x_1 x_2} + \nabla_k^2 F + L(w, F) - q &= 0.
 \end{aligned} \tag{4.25}$$

By solving (4.6) with respect to  $\varepsilon_{ik}$ , one gets

$$\begin{aligned}
 \varepsilon_{11} &= AF_{,x_2 x_2} + A*F_{,x_1 x_1} + Bw_{,x_1 x_1} = B* w_{,x_2 x_2}, \\
 \varepsilon_{22} &= A*F_{,x_2 x_2} + AF_{,x_1 x_1} + B* w_{,x_1 x_1} + Bw_{,x_2 x_2}, & \varepsilon_{12} &= -4A_2 F_{,x_1 x_2} + 4B_2 w_{,x_1 x_2},
 \end{aligned} \tag{4.26}$$

where

$$\begin{aligned}
 A &= A_1 + A_2, & A* &= A_1 - A_2, & B &= B_1 + B_2, & B* &= B_1 - B_2, \\
 A_1 &= \frac{0.5}{c_{00} + c_{10}}, & A_2 &= \frac{0.5}{c_{00} - c_{10}}, \\
 B_1 &= (c_{01} + c_{11})A_1, & B_2 &= (c_{01} - c_{11})A_2.
 \end{aligned} \tag{4.27}$$

Taking into account (4.27) in (4.9), one obtains

$$\begin{aligned} M_{11} &= BF_{,x_2x_2} + B^*F_{,x_1x_1} - Cw_{,x_1x_1} - C^*w_{,x_2x_2}, \\ M_{22} &= B^*F_{,x_2x_2} + BF_{,x_1x_1} - C^*w_{,x_1x_1} - Cw_{,x_2x_2}, \\ M_{12} &= -2B_2F_{,x_1x_2} - (C - C^*)w_{,x_1x_2}, \end{aligned} \quad (4.28)$$

where

$$\begin{aligned} C &= c_{02} - B_1^* - B_2^*, & C^* &= c_{12} - B_1^* - B_2^*, \\ B_1^* &= (c_{01} + c_{11})B_1, & B_2^* &= (c_{01} - c_{11})B_2. \end{aligned} \quad (4.29)$$

Equations (4.25) with an account of (4.27), (4.28) read

$$\begin{aligned} & (A^*F_{,x_1x_1} + AF_{,x_2x_2} + Bw_{,x_1x_1} + B^*w_{,x_2x_2})_{,x_2x_2} \\ & + (AF_{,x_1x_1} + A^*F_{,x_2x_2} + B^*w_{,x_1x_1} + Bw_{,x_2x_2})_{,x_1x_1} \\ & - (-4A_2F_{,x_1x_2} + 4B_2w_{,x_1x_2})_{,x_1x_2} + \nabla_k^2 w + 0.5L(w, w) = 0; \\ & (BF_{,x_1x_1} + B^*F_{,x_2x_2} - C^*w_{,x_1x_1} - Cw_{,x_2x_2})_{,x_2x_2} \\ & + (B^*F_{,x_1x_1} + BF_{,x_2x_2} - Cw_{,x_1x_1} - C^*w_{,x_2x_2})_{,x_1x_1} \\ & - 2(-2B_2F_{,x_1x_2} - (C - C^*)w_{,x_1x_2})_{,x_1x_2} \\ & + \nabla_k^2 F + L(w, F) - q = 0. \end{aligned} \quad (4.30)$$

In what follows, the algorithm for solving the stability problem of physically nonlinear flexible shallow shells with mixed boundary conditions and without ribs is discussed.

*Step 1.* In the space occupied by the shell, a rectangular mesh with equal step  $\{x_{1i}, x_{2j}, x_{3k}\}$  is introduced. Assuming in the beginning that the shell is unloaded, modulus  $G_{ijk}$  of the shell nodes is introduced.

*Step 2.* At the nodes of mesh surface  $x_3 = 0$ , coefficients  $C_{\alpha\beta}$  (4.11) are computed. Then, substituting the partial derivatives by difference expressions with error  $Q(h_{x_1}^2 + h_{x_2}^2)$ , instead of the system of equations (4.30), the system of nonlinear algebraic equations with respect to  $w_{ij}$ ,  $F_{ij}$ , that is, deflection and stress functions formulated at nodes of the mean surface, is obtained. Boundary conditions are considered while out-contour nodes are being eliminated from the equations system. If the mesh node overlaps with the point where the boundary conditions change, for instance, in the case of hinge clamping, then the clamping is understood as the value of derivative with respect to a normal solution to a system of algebraic equations for a given load  $q(x_1x_2)$  defined with the help of the Newton or differentiation along the parameter methods [7], and then the elimination method is applied.

*Step 3.* New values of  $G_{ijk}$  and  $\mu_{ijk}$  are computed at the mesh nodes in accordance with the theory of small elastic-plastic deformation [1, 4]. Modulus of elongation and modulus of volume deformation  $K$  are found using formulas of the theory of small elastic-plastic deformation

$$E = \frac{9KG}{3K + G}, \quad \mu = \frac{1}{2} \frac{3K - 2G}{3K + G}. \quad (4.31)$$

Notice that in this theory, it is assumed that  $K$  does not depend on the deformable state at a point since in the case of a body which is nonhomogeneous before deformation,  $K = K(x_1, x_2, x_3)$ , and for a body homogeneous before deformation,  $K = K_0 = \text{const}$ .

The shear modulus is defined through the formula

$$G = \frac{1}{3} \frac{\sigma_i(e_i)}{e_i(x)}, \quad (4.32)$$

and  $G$  is called the cutting modulus. In order to compute the shell, dependence  $\sigma_i(e_i)$  should be explicitly given (intensity of stress  $\sigma_i$  versus strain  $(e_i)$ ).

Some of the diagrams  $\sigma_i(e_i)$  are reported in [4]. Then the procedure goes back to [Step 2](#).

The computations are repeated until the obtained solution overlaps with the previous one keeping the assumed accuracy.

*Step 4.* The load  $q = q + h_q$  is increased and the computations start at [Step 2](#).

As a result, dependence  $q(w)$  yielding the critical values is obtained.

On a basis of the introduced algorithm, the program in Fortran has been developed.

As an example, a square shell made of aluminium and with parameters  $k_1 = k_2 = 18$ ,  $\lambda = a/h = 50$ , and  $q = \text{const}$  [4] has been considered.

We apply the Mises flow condition ( $\sigma_i = \sigma_s$ ) and the dependence  $\sigma_i(e_i)$  taken in the form

$$\sigma_i = \sigma_s \left( 1 - \exp \frac{-e_i}{e_s} \right), \quad (4.33)$$

where  $e_s = \sigma_s/3\sigma_0$  is the artificial intensity of the flow deformation.

The problem is considered symmetric with respect to  $x_1, x_2$  for four types of the boundary conditions, namely,

- (1) hinged support on the contour;
- (2) hinged support in an angle, and clamping in the middle of the side over the interval  $1/4a$ ;
- (3) clamping in an angle, and hinged support in the middle of the side over the interval  $1/4a$ ;
- (4) clamping along a contour.

The stress function on the side  $x_2 = 0$  is  $F_{,x_1x_1} = 0$ ,  $F_{,x_1x_2} = 0$ , whereas for the deflection,  $w = 0$ ,  $w_{,x_2} = 0$  or  $M_{22} = 0$ .

Consider a few computational results. In [Figure 4.2](#), dependencies of load versus center deflection represented by curves 1–4 corresponding to the considered boundary conditions are reported. Analyzing the behavior of the curves, one may conclude that the value of the upper critical load depends essentially on the position of the clamped part. [Figure 4.3](#) shows influence of the boundary conditions on the distribution of plasticity zones on the upper shell surface  $x_3 = -h/2$  for the loads corresponding to deflection  $w = 0, 4$  in the center. The largely developed plasticity zones are achieved for hinged support. Owing to clamping of the shell point in the middle of its side, the plastic zones are sharply decreased and their position is subject to change. The minimal areas of the plasticity zones are achieved in the case of clamping along the whole length.



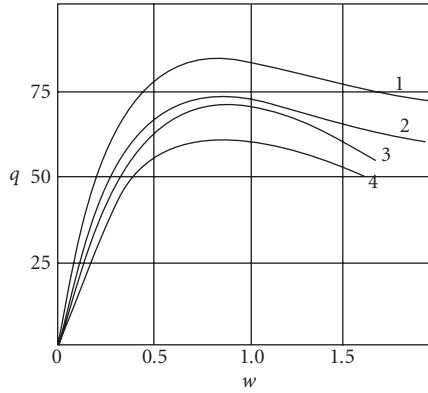


Figure 4.2. Load versus shell center deflection (see text for more details).

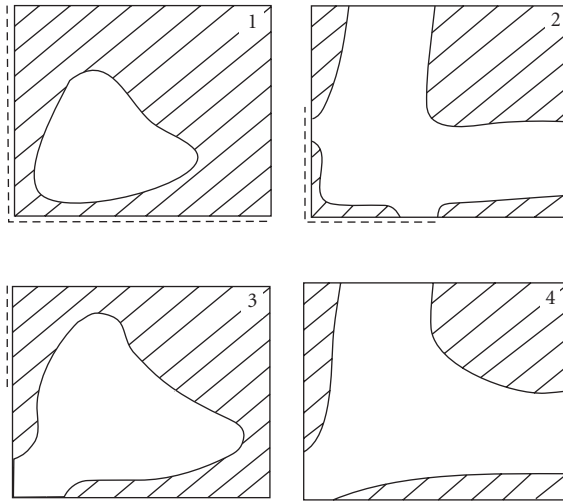


Figure 4.3. Influence of boundary conditions on plasticity zone distribution (see text for more details).

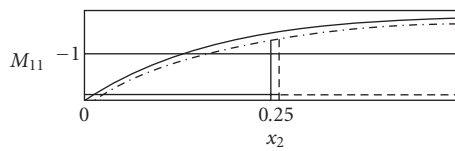


Figure 4.4. Moment  $M_{11}$  versus  $x_2$  on shell side  $x_1 = 0$  (see text).

Figure 4.4 displays graphs of moments  $M$  on the side  $x_1 = 0$  for the following cases: 2—solid curve; 3—dashed curve; 4—dash-dotted curve.

Deflection in the shell center achieved in all cases is  $w = 0.4$ . In Figures 4.2 and 4.3, digits 1–4 should be considered simultaneously. For the given computational model, contrary to [8], a jump of the moments is observed, which is close to real observations.

The results of Section 4 make the elimination method for equations effective and also for solving the complicated PDEs like nonlinear equations governing elastic-plastic problems of shells with finite deflections. Here, on each loading step, one has to solve many times a large SLAE, which, without any doubt, exerts significant impact on the chosen computational algorithm.

## References

- [1] J. Awrejcewicz and V. A. Krysko, *Nonclassical Thermoelastic Problems in Nonlinear Dynamics of Shells*, Scientific Computation, Springer-Verlag, Berlin, 2003.
- [2] V. I. Feodos'ev, *On a method of solution of the nonlinear problems of stability of deformable systems*, J. Appl. Math. Mech. **27** (1963), 392–404.
- [3] G. E. Forsythe and C. B. Moler, *Computer Solution of Linear Algebraic Systems*, Prentice-Hall, New Jersey, 1967.
- [4] V. A. Krysko, *Nonlinear Statics and Dynamics of Non-Homogeneous Shells*, Saratov State University Press, Saratov, 1976.
- [5] A. A. Samarskii and V. B. Andreev, *Difference Methods for Elliptic Equations*, Nauka, Moscow, 1976.
- [6] A. S. Saninskiy, *Direct Method of Solution for a System of Linear Algebraic Equations with Calculated Matrix*, Saratov State University Press, Saratov, 1877.
- [7] V. I. Shalashvili and E. B. Kusnetsov, *Application of Solution along the Parameter Method and on the Best Parametrization*, Editorial URSS, Moscow, 1999.
- [8] L. F. Vakhlaeva and V. A. Krysko, *Stability of flexible shallow rectangular shells with various boundary conditions on their sides*, Building and Architecture. IV, VUZ Press, Kazan, 1984, pp. 21–25.

Jan Awrejcewicz: Department of Automatics and Biomechanics, Technical University of Lodz, 90-924 Lodz, Poland

*E-mail address:* [awrejcew@mail.p.lodz.pl](mailto:awrejcew@mail.p.lodz.pl)

Vadim A. Krysko: Department of Mathematics, Saratov state University, 410054 Saratov, Russia

*E-mail address:* [tek@sun.ru](mailto:tek@sun.ru)

Anton V. Krysko: Department of Mathematics, Saratov state University, 410054 Saratov, Russia