Emil Humhal One iterative method concerning the solution of Dirichlet's problem

Aplikace matematiky, Vol. 21 (1976), No. 2, 120-135

Persistent URL: http://dml.cz/dmlcz/103630

# Terms of use:

© Institute of Mathematics AS CR, 1976

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

# ONE ITERATIVE METHOD CONCERNING THE SOLUTION OF DIRICHLET'S PROBLEM

#### Emil Humhal

(Received July 17, 1974)

### 1. INTRODUCTION

In the article, we shall study the solution by the method of finite differences of the equation

(1) 
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - \Sigma(x, y) u = \chi(x, y)$$

on the square  $\langle 0, 1 \rangle \times \langle 0, 1 \rangle$  with the boundary condition  $u(x, y) = \psi(x, y)$  on the boundary. The generalization of the method to an area whose boundary is formed by segments parallel to the axes x and y presents no difficulties. In the whole work the differential considerations are carried out in a convenient finite difference transcription.

The basic aim of our method is to produce the solution of the original problem by iterating the equation of the form

$$\boldsymbol{L}_{xy}(u) = \boldsymbol{F}(u)$$

obtained by a suitable arrangement of equation (1) with the same boundary condition, where  $L_{xy}$  and F are appropriate operators. We require the operator  $L_{xy}$  to be decomposable as a product of two operators  $L_x$  and  $L_y$  each of which depends on a single variable x or y only, i.e.  $L_{xy} = L_x L_y$ . This leads to an equivalent equation to (1) of the form

$$\boldsymbol{L}_{\boldsymbol{x}}(\boldsymbol{L}_{\boldsymbol{y}}\boldsymbol{u}) = \boldsymbol{F}(\boldsymbol{u}) \, .$$

To find u an iterative process is proposed according to which u is a limit of a sequence  $\{u^{(k)}\}$  defined by solving the system

$$\mathbf{L}_{x}(\mathbf{L}_{y}u^{(k)}) = \mathbf{F}(u^{(k-1)}), \quad u^{(k)}(x, y) = \psi(x, y)$$

120

.

on the boundary (k = 0, 1, 2, ...). The k-th iteration will be then found as the solution of two boundary-value problems of one variable

$$L_x \varphi^{(k)} = F(u^{(k-1)}), \quad \varphi^{(k)}(x, y) = L_y \psi(x, y) \text{ for } x = 0 \text{ and } x = 1$$

and

$$L_y u^{(k)} = \varphi^{(k)}$$
,  $u^{(k)}(x, y) = \psi(x, y)$  for  $y = 0$  and  $y = 1$ 

In order to make the explanation easier we shall analyse separately the case when the coefficient  $\Sigma$  is constant.

### 2. THE CASE OF CONSTANT COEFFICIENT $\varSigma$

Let us assume in this section that  $\Sigma$  is a positive constant. We choose a positive integer *n* and on the square we form the square net of mesh spacing h = 1/n. When speaking of the point (i, j), we mean the point with the coordinates [ih, jh]. The symbol  $u_{ij}$  denotes both the exact and the approximate values of the function *u* at the point (i, j). It will be evident from the context whether it denotes the exact or the approximate value. The same notation will be used for other functions. We shall form the iterative sequence of vectors  $\mathbf{u}^{(k)}$  with components  $u_{ij}^{(k)}$  (i, j = 0, 1, ..., n) according to the symbolical formula

(2) 
$$\frac{\partial^4 u^{(k)}}{\partial x^2 \partial y^2} - \Sigma \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) u^{(k)} + \Sigma^2 u^{(k)} = -\Sigma \chi + \frac{\partial^4 u^{(k-1)}}{\partial x^2 \partial y^2},$$

 $u^{(k)}(x, y) = \psi(x, y)$  on the boundary.

The formula (2) could be rewritten as

(3) 
$$\left(\frac{\partial^2}{\partial x^2} - \Sigma I\right) \left(\frac{\partial^2}{\partial y^2} - \Sigma I\right) u^{(k)} = -\Sigma \chi + \frac{\partial^4 u^{(k-1)}}{\partial x^2 \partial y^2},$$

where I denotes the identity operator. When replacing the derivatives by the current difference formulas, (2) will change to the system

$$(2') \quad (2 + \Sigma h^2)^2 u_{ij}^{(k)} - (2 + \Sigma h^2) (u_{i+1,j}^{(k)} + u_{i-1,j}^{(k)} + u_{i,j+1}^{(k)} + u_{i,j-1}^{(k)}) + + (u_{i+1,j+1}^{(k)} + u_{i+1,j-1}^{(k)} + u_{i-1,j+1}^{(k)} + u_{i-1,j-1}^{(k)}) = 4u_{ij}^{(k-1)} - - 2(u_{i+1,j}^{(k-1)} + u_{i-1,j}^{(k-1)} + u_{i,j+1}^{(k-1)} + u_{i,j-1}^{(k-1)}) + + (u_{i+1,j+1}^{(k-1)} + u_{i+1,j-1}^{(k-1)} + u_{i-1,j+1}^{(k-1)} + u_{i-1,j-1}^{(k-1)}) - \Sigma h^4 \chi_{ij}$$

for every inner point (i, j) of the net while we put

$$u_{ij}^{(k)} = \psi_{ij}$$

if i or j is equal to 0 or n.

It is obvious that in the case of convergence the limit vector fulfils the current five point formula equations. Denoting the known right-hand side in (2') by  $\mu_{ij}^{(k-1)}$  we obtain the vector  $\mathbf{u}^{(k)}$  through the solution of two systems of equations (4) and (5)

$$\varphi_{0j}^{(k)} = \psi_{0,j+1} - (2 + \Sigma h^2) \psi_{0j} + \psi_{0,j-1}$$

$$\varphi_{0j}^{(k)} - (2 + \Sigma h^2) \varphi_{1j}^{(k)} + \varphi_{2j}^{(k)} = \mu_{1j}^{(k-1)}$$
(4)
$$\varphi_{1j}^{(k)} - (2 + \Sigma h^2) \varphi_{2j}^{(k)} + \varphi_{3j}^{(k)} = \mu_{2j}^{(k-1)}$$

$$\vdots$$

$$\varphi_{n-2,j}^{(k)} - (2 + \Sigma h^2) \varphi_{n-1,j}^{(k)} + \varphi_{nj}^{(k)} = \mu_{n-1,j}^{(k-1)}$$

$$\varphi_{nj}^{(k)} = \psi_{n,j+1} - (2 + \Sigma h^2) \psi_{nj} + \psi_{n,j-1}$$

for all j = 1, 2, ..., n - 1,

$$u_{i0}^{(k)} = \psi_{i0}$$

$$u_{i0}^{(k)} - (2 + \Sigma h^{2}) u_{i1}^{(k)} + u_{i2}^{(k)} = \varphi_{i1}^{(k)}$$

$$u_{i1}^{(k)} - (2 + \Sigma h^{2}) u_{i2}^{(k)} + u_{i3}^{(k)} = \varphi_{i2}^{(k)}$$

$$\vdots$$

$$u_{i,n-2}^{(k)} - (2 + \Sigma h^{2}) u_{i,n-1}^{(k)} + u_{in}^{(k)} = \varphi_{i,n-1}^{(k)}$$

$$u_{i,n-2}^{(k)} = \psi_{in}$$

for all i = 1, 2, ..., n - 1.\*

Now we shall deal with the convergence of the iterative sequence. Certainly it is sufficient to prove convergence only for the vector  $\mathbf{u}^{(k)}$ , which differs from the original one by omitting the components corresponding to the boundary points. Let us denote

$$\mathbf{u}^{(k)} = \left[u_{1,1}^{(k)}, u_{1,2}^{(k)}, \dots, u_{1,n-1}^{(k)}, u_{2,1}^{(k)}, u_{2,2}^{(k)}, \dots, u_{2,n-1}^{(k)}, \dots, u_{n-1,1}^{(k)}, u_{n-1,2}^{(k)}, \dots, u_{n-1,n-1}^{(k)}\right]^{\mathsf{T}}.$$

\*) It is worth considering whether it would not be possible to use analogously the fact that the operator

$$\frac{\partial^4}{\partial x^2 \partial y^2} - \Sigma \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + \Sigma^2 \mathbf{I}$$

admits also the decomposition

$$\left(\frac{\partial^2}{\partial x \,\partial y} + \sqrt{(\Sigma)} \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) + \Sigma I\right) \left(\frac{\partial^2}{\partial x \,\partial y} - \sqrt{(\Sigma)} \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) + \Sigma I\right)$$
$$\left(\frac{\partial}{\partial x} + \sqrt{(\Sigma)} I\right) \left(\frac{\partial}{\partial y} + \sqrt{(\Sigma)} I\right) \left(\frac{\partial}{\partial x} - \sqrt{(\Sigma)} I\right) \left(\frac{\partial}{\partial y} - \sqrt{(\Sigma)} I\right)$$

or

respectively, all provided the derivatives commute. While the original decomposition leads to a boundary problem, these ones lead to a Cauchy problem.

Further let us denote by **A** a block diagonal matrix with n - 1 diagonal blocks (everywhere in this section the blocks will be of the type  $(n - 1) \times (n - 1)$ )

Let us denote  $\mathbf{A}(h) = \mathbf{A} + \Sigma h^2 \mathbf{I}$ . Obviously  $\mathbf{A}(h) \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}(h)$  is valid and both matrices are symmetric and positive definite.\*)

Let us denote by **P** the permutation matrix, written blockwise as

(6) 
$$\mathbf{P} = \begin{vmatrix} \mathbf{E}_{11} & \mathbf{E}_{21} & \dots & \mathbf{E}_{n-1,1} \\ \mathbf{E}_{12} & \mathbf{E}_{22} & \dots & \mathbf{E}_{n-1,2} \\ \dots & \dots & \dots & \dots \\ \mathbf{E}_{1,n-1} & \mathbf{E}_{2,n-1} & \dots & \mathbf{E}_{n-1,n-1} \end{vmatrix}$$

where  $\mathbf{E}_{ij}$  is the square matrix, with all elements zero except for that at the place (i, j) which is equal to one. We can see very easily that  $\mathbf{P} = \mathbf{P}^{\mathsf{T}}$  and  $\mathbf{P}^{\mathsf{T}}\mathbf{P} = \mathbf{I}$ . It is evident that

When we arrange system (2') so that the known quantities corresponding to the boundary points are on the right-hand side then the system (2') is transformed to the matrix form

$$\mathbf{B}(h) \mathbf{u}^{(k)} = \mathbf{B}(0) \mathbf{u}^{(k-1)} + \mathbf{f},$$

where **f** is a vector which does not depend on k and **B**(h) is the block matrix

$$\mathbf{B}(h) = \begin{bmatrix} \mathbf{B}_{1}(h) & \mathbf{B}_{2}(h) & & \\ \mathbf{B}_{2}(h) & \mathbf{B}_{1}(h) & \mathbf{B}_{2}(h) & \\ & \mathbf{B}_{2}(h) & \mathbf{B}_{1}(h) & \mathbf{B}_{2}(h) & \\ & \ddots & \ddots & \\ & & & \mathbf{B}_{2}(h) \\ & & \mathbf{B}_{2}(h) & \mathbf{B}_{1}(h) \end{bmatrix}$$

\*) As far as we consider positive definite matrices we suppose automatically that they are symmetric.

where

$$\mathbf{B}_{1}(h) = \begin{bmatrix} (2 + \Sigma h^{2})^{2} & -(2 + \Sigma h^{2}) \\ -(2 + \Sigma h^{2}) & (2 + \Sigma h^{2})^{2} & -(2 + \Sigma h^{2}) \\ & \ddots & \ddots & & \\ & & -(2 + \Sigma h^{2}) \\ & & -(2 + \Sigma h^{2}) & (2 + \Sigma h^{2})^{2} \end{bmatrix} = \\ = (2 + \Sigma h^{2}) (\mathbf{A}_{1} + \Sigma h^{2} \mathbf{I}) \text{ and } \mathbf{B}_{2}(h) = -(\mathbf{A}_{1} + \Sigma h^{2} \mathbf{I}).$$

Consequently, the convergence of the method is defined by the spectral radius of the matrix  $K_1 = B^{-1}(h) B(0) = A^{-1}(h) PA^{-1}(h) APA$ . We may write  $K_1 = A^{-1}K_2A$ , where  $K_2 = AA^{-1}(h) PA^{-1}(h) AP$ . We shall prove that all eigenvalues of the matrix  $K_2^{-1}$  are real and greater than 1. It holds  $K_2^{-1} = PA^{-1} A(h) P A(h) A^{-1}$ . The relation  $A^{-1}A(h) = A(h) A^{-1} = I + \Sigma h^2 A^{-1}$  implies that the matrix  $A^{-1} A(h) = S^2$ . Then ought to be  $K_2^{-1} = S^{-1}K_3S$ , where  $K_3 = SPS^2PS$ . Let us establish a lower estimate of eigenvalues of the symmetric matrix  $K_3$ :

$$\frac{(\mathbf{K}_{3}\mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})} = \frac{(\mathbf{SPS}^{2}\mathbf{PS}\mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})} = \frac{(\mathbf{SP}[\mathbf{I} + \Sigma h^{2}\mathbf{A}^{-1}] \mathbf{PS}\mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})} =$$
$$= \frac{([\mathbf{S}^{2} + \Sigma h^{2}\mathbf{SPA}^{-1}\mathbf{PS}] \mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})} = \frac{([\mathbf{I} + \Sigma h^{2}\mathbf{A}^{-1} + \Sigma h^{2}\mathbf{SPA}^{-1}\mathbf{PS}] \mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})}$$
$$= \frac{(\mathbf{x},\mathbf{x}) + \Sigma h^{2}\{(\mathbf{A}^{-1}\mathbf{x},\mathbf{x}) + (\mathbf{A}^{-1}\mathbf{PS}\mathbf{x},\mathbf{PS}\mathbf{x})\}}{(\mathbf{x},\mathbf{x})} > 1,$$

for an arbitrary vector  $\mathbf{x} \neq 0$ .

### 3. The case of non-constant $\varSigma$

Now we shall deal with the case when the coefficient  $\Sigma$  in (1) is a function of the variables x and y. First let us suppose that the functions  $\Sigma$  and  $\chi$  as well as the boundary condition  $\psi$  are continuous. Moreover, we assume in the whole paper, that the function  $\Sigma$  is non-negative. Since we shall keep the notation of the previous section, the equations for the current five points approximation scheme have the form

(7) 
$$(\mathbf{A} + \mathbf{P}\mathbf{A}\mathbf{P})\mathbf{u} + h^2\Sigma\mathbf{u} = \mathbf{f},$$

where f is a vector given by the right-hand side of the differential equation and by the boundary conditions while  $\Sigma$  is a diagonal matrix with non-negative elements equal

to the value of the function  $\Sigma$  at the inner points of the net. We shall form the iterative sequence of the vectors  $\mathbf{u}^{(k)}$  according to the formula

(8) 
$$(\mathbf{A} + \omega h^2 \mathbf{I}) (\mathbf{P} \mathbf{A} \mathbf{P} + \omega h^2 \mathbf{I}) \mathbf{u}^{(k)} = \{\mathbf{A} \mathbf{P} \mathbf{A} \mathbf{P} + \omega h^4 (\omega \mathbf{I} - \Sigma)\} \mathbf{u}^{(k-1)} + \omega h^2 \mathbf{f},$$

where  $\omega \neq 0$  is a constant, which, for example, can be taken as suitable approximation of the function  $\Sigma$ . The formula corresponds to the transcription of the differential equation (1) of the form

(9) 
$$\left(\frac{\partial^2}{\partial x^2} - \omega I\right) \left(\frac{\partial^2}{\partial y^2} - \omega I\right) u = \frac{\partial^4 u}{\partial x^2 \partial y^2} + \omega^2 u - \omega \Sigma u - \omega \chi .$$

Now we shall prove the convergence of the iterative sequence (obviously in dependence on the choice of  $\omega$ ). We shall use the following lemma.

*Lemma*: Let the matrices **A** and **B** be either both positive or both negative definite and commuting. Then the matrix AB = BA is positive definite.

*Proof*: The matrix **AB** is symmetric because  $(\mathbf{AB})^{\mathsf{T}} = \mathbf{BA} = \mathbf{AB}$  and therefore it is sufficient to prove that all eigenvalues of **AB** are positive. Let us write **A** in the form  $\mathbf{A} = \mathbf{S}^2$  provided **A** is positive definite and  $\mathbf{A} = -\mathbf{S}^2$  for **A** negative definite, where **S** is a positive definite matrix. Then  $\mathbf{AB} = \mathbf{S}(\mathbf{SBS})\mathbf{S}^{-1}$  or  $\mathbf{AB} = \mathbf{S}(\mathbf{S}(-\mathbf{B})\mathbf{S})\mathbf{S}^{-1}$  respectively. Therefore all eigenvalues of **AB** are positive.

It follows from the lemma that the matrix on the left-hand side of (8) is positive definite as far as  $\omega h^2$  does not belong to the interval containing all eigenvalues of the matrix  $-\mathbf{A}$ . Let  $\omega$  fulfill the condition. The matrix on the right-hand side of (8) is symmetric. The convergence is determined by the eigenvalues of the matrix

(10) 
$$\mathbf{C} = \left[ (\mathbf{A} + \omega h^2 \mathbf{I}) (\mathbf{P} \mathbf{A} \mathbf{P} + \omega h^2 \mathbf{I}) \right]^{-1} \left\{ \mathbf{A} \mathbf{P} \mathbf{A} \mathbf{P} + \omega h^4 (\omega \mathbf{I} - \boldsymbol{\Sigma}) \right\}.$$

It is possible to write  $[(\mathbf{A} + \omega h^2 \mathbf{I}) (\mathbf{PAP} + \omega h^2 \mathbf{I})]^{-1} = \mathbf{S}^2$ , where **S** is positive definite and so  $\mathbf{C} = \mathbf{SQS}^{-1}$ , where

(11) 
$$\mathbf{Q} = \mathbf{S} \cdot \{\mathbf{APAP} + \omega h^4 (\omega \mathbf{I} - \Sigma)\} \mathbf{S}$$

is a symmetric matrix. It holds

(12) 
$$\mathbf{Q} = \mathbf{S} \cdot \{ (\mathbf{A} + \omega h^2 \mathbf{I}) (\mathbf{P} \mathbf{A} \mathbf{P} + \omega h^2 \mathbf{I}) - \omega h^2 \mathbf{A} - \omega h^2 \mathbf{P} \mathbf{A} \mathbf{P} - \omega h^4 \Sigma \} \mathbf{S} =$$
$$= \mathbf{I} - \omega h^2 \mathbf{S} \mathbf{A} \mathbf{S} - \omega h^2 \mathbf{S} \mathbf{P} \mathbf{A} \mathbf{P} \mathbf{S} - \omega h^4 \mathbf{S} \Sigma \mathbf{S} .$$

The condition of the convergence is fulfilled, when all vectors  $\mathbf{u} \neq 0$  satisfy

(13) 
$$0 < \frac{\omega h^2 (\mathbf{S}(\mathbf{A} + \mathbf{P}\mathbf{A}\mathbf{P}) \mathbf{S}\mathbf{u}, \mathbf{u}) + \omega h^4 (\mathbf{S}\boldsymbol{\Sigma}\mathbf{S}\mathbf{u}, \mathbf{u})}{(\mathbf{u}, \mathbf{u})} < 2$$

that is

$$0 < \frac{\omega h^2((\mathbf{A} + \mathbf{PAP}) \mathbf{Su}, \mathbf{Su}) + \omega h^4(\Sigma \mathbf{Su}, \mathbf{Su})}{(\mathbf{u}, \mathbf{u})} < 2$$

If we denote  $\mathbf{v} = \mathbf{S}\mathbf{u}$ , we require that all  $\mathbf{v} \neq 0$  satisfy

$$0 < \frac{\omega h^2((\mathbf{A} + \mathbf{PAP})\mathbf{v}, \mathbf{v}) + \omega h^4(\boldsymbol{\Sigma}\mathbf{v}, \mathbf{v})}{(\mathbf{S}^{-1}\mathbf{v}, \mathbf{S}^{-1}\mathbf{v})} < 2,$$

that is

$$0 < \frac{\omega h^2((\mathbf{A} + \mathbf{PAP})\mathbf{v}, \mathbf{v}) + \omega h^4(\Sigma \mathbf{v}, \mathbf{v})}{(\mathbf{APAPv}, \mathbf{v}) + \omega h^2((\mathbf{A} + \mathbf{PAP})\mathbf{v}, \mathbf{v}) + \omega^2 h^4(\mathbf{v}, \mathbf{v})} < 2,$$

which is fulfilled for example for  $\omega \ge \max_{\substack{0 \le x \le 1 \\ 0 \le y \le 1}} \Sigma(x, y)$ .

Further we shall study the discontinuous case. Let us suppose that the points of discontinuity of the functions  $\Sigma$  and  $\chi$  are situated in such a way that the square  $\langle 0, 1 \rangle \times \langle 0, 1 \rangle$  is the closure of the union of disjoint rectangles such that the functions  $\Sigma$  and  $\chi$  are continuous inside any rectangle, the sides of these rectangles lay on straight lines of the net and any two neighbouring rectangles have a common side of the same length.

Now we form the system of difference equations. For this purpose we divide the straight lines of the net (excluding the boundary ones) into the straight lines of the first and of the second sort. We include all straight lines on which the sides of the just mentioned rectangles lay and, if we want to, also some other straight lines into the family of straight lines of the second sort. All the other straight lines are called the straight lines of the first sort. Let us suppose that the net is so dense and the division made in such a way that neither two straight lines of the second sort nor the straight line of the second sort and the boundary one neighbour anywhere.

For an inner mesh-point (i, j) which is the intersection of two straight lines of the first sort we require the following equation to be fulfilled:

(14) 
$$4u_{ij} - u_{i+1,j} - u_{i-1,j} - u_{i,j+1} - u_{i,j-1} + h^2 \Sigma_{ij} u_{ij} = -h^2 \chi_{ij}$$

For any boundary mesh-point which lies on a straight line of the first sort we put

$$(15) u_{ij} = \psi_{ij} \,.$$

The same should hold for the vertices of the square. For the mesh-points (i, j) (including boundary ones) lying on a straight line of the second sort parallel with the axis y we require the following equation to be fulfilled:

(16) 
$$u_{i-2,j} - 4u_{i-1,j} + 6u_{ij} - 4u_{i+1,j} + u_{i+2,j} = 0.$$

For mesh-points lying on a straight-line of the second sort parallel with the axis x, which are points of intersection with straight lines of the first sort or which are boundary mesh-points, we require the following equation to be fulfilled:

(17) 
$$u_{i,j-2} - 4u_{i,j-1} + 6u_{ij} - 4u_{i,j+1} + u_{i,j+2} = 0.$$

For mesh-points lying on these straight lines which are points of intersection with straight lines of the second sort the fulfilment of (16) was already required. The equation (17) is, however, fulfilled automatically, which we shall prove. Let us suppose that a given straight line of the second sort parallel with the axis x includes the mesh-points (i, j) for i = 0, 1, ..., n (that is, i = 0, n corresponds to the boundary mesh-points). Let us denote

$$\varphi_i = u_{i,j-2} - 4u_{i,j-1} + 6u_{ij} - 4u_{i,j+1} + u_{i,j+2}.$$

It follows from (17) that  $\varphi_i = 0$  for indices *i* corresponding to the intersection of our straight line and a straight line of the first sort or the boundary. Particularly,  $\varphi_0 = \varphi_1 = \varphi_{n-1} = \varphi_n = 0$ . As far as *i* is the index corresponding to the intersection of the straight line of the second sort, our conditions imply  $\varphi_{i-1} = \varphi_{i+1} = 0$ . From (16) it is evident that

$$\varphi_{i-2} + 6\varphi_i + \varphi_{i+2} = 0$$

for every point of such type where  $\varphi_{i-2}$  or  $\varphi_{i+2}$  could be zero. Hence  $\varphi_i$  for the examined indices fulfil a homogeneous system of linear algebraic equations with a regular matrix and therefore all  $\varphi_i$  are zero. This guarantees that our system is the same when numbering the mesh-points first in the direction of the axis x and then in the direction of the axis y as for the reverse order.

The equations (16) and (17) are difference transcriptions of the condition of continuity of  $\partial u/\partial x$  and  $\partial u/\partial y$ , respectively. We cannot use simpler transcriptions of this condition

(16') 
$$u_{i-1,j} - 2u_{ij} + u_{i+1,j} = 0,$$

(17') 
$$u_{i,j-1} - 2u_{ij} + u_{i,j+1} = 0$$

because their precision is not in accordance with (14). How to solve the obtained system by iterations will be shown for a special choice of the straight lines of the first and the second sorts.

We shall solve again the equation (1) on the square  $\langle 0, 1 \rangle \times \langle 0, 1 \rangle$ . We choose a positive integer *n* and denote h = 1/(2n). The straight lines x = ih (i = 1, 3, 5, ..., 2n - 1) and y = jh (j = 1, 3, 5, ..., 2n - 1) will be considered the straight lines of the first sort; the straight lines x = ih (i = 2, 4, ..., 2n - 2) and y = jh(j = 2, 4, ..., 2n - 2) the straight lines of the second sort. In figure 1, we denote intersections of two straight lines of the first sort by a little cross and the intersections of straight lines of the second sort with straight lines of the first sort by a circle. We shall then speak about crossed and circled points. From the preceding argument it follows that we require (14) to be fulfilled at the crossed points and (16) at the circled ones provided the points lie on a straight line of the first sort (or a boundary one) parallel with the axis x; otherwise we require (17) to be fulfilled.



Fig. 1.

At the remaining boundary points we require the boundary condition and at the remaining inner points we require (16) (and consequently also (17)). On the set of vectors  $\mathbf{v}$  with  $n^2$  components

$$\mathbf{v} = \begin{bmatrix} u_{11}, u_{13}, \dots, u_{1,2n-1}, u_{31}, u_{33}, \dots, u_{3,2n-1}, \dots, u_{2n-1,1}, u_{2n-1,3}, \dots, u_{2n-1,2n-1} \end{bmatrix}^{\mathsf{T}}$$

corresponding to the vectors of values of function u at the crossed points, we define the operator  $\mathbf{L}_{\omega}$  (in the correspondence with the operator  $(\partial^2/\partial x^2 - \omega \mathbf{I}) \cdot (\partial^2/\partial y^2 - \omega \mathbf{I})$ ) in the following way. Let us define for a moment auxiliary values  $u_{ij}$  for the indices (i, j) corresponding to other points than the crossed ones from the conditions (16) or (17) or from the boundary condition respectively in the way required in our system of difference equations. Then we define

$$\begin{bmatrix} \mathbf{L}_{\omega} \mathbf{v} \end{bmatrix}_{ij} = (u_{i+1,j-1} - (2 + \omega h^2) u_{i+1,j} + u_{i+1,j+1}) - (2 + \omega h^2) (u_{i,j-1} - (2 + \omega h^2) u_{ij} + u_{i,j+1}) + (u_{i-1,j-1} - (2 + \omega h^2) u_{i-1,j} + u_{i-1,j+1})$$

for the relevant indices (i, j).

If we form the iterative sequence according to

(18) 
$$\mathbf{L}_{\omega}\mathbf{v}^{(k)} = \mathbf{L}_{0}\mathbf{v}^{(k-1)} + \omega h^{4}(\omega \mathbf{I} - \Sigma)\mathbf{v}^{(k-1)} - \omega h^{4}\mathbf{\chi}$$

where  $\Sigma$  is the diagonal matrix with the elements  $\Sigma_{ij}$  and  $\chi$  is the vector with the elements  $\chi_{ij}$  with indices corresponding to the crossed points, then in the case of

$$\begin{aligned} & (20) \\ & u_{2i-1,0}^{(0)} - (2 + \omega h^2) u_{2i-1,1}^{(0)} + u_{2i-1,2}^{(0)} = \phi_{2i-1,1} \\ & u_{2i-1,2}^{(0)} - 4u_{2i-1,1}^{(0)} + 6u_{2i-1,2}^{(0)} - 4u_{2i-1,3}^{(0)} + u_{2i-1,3}^{(0)} + u_{2i-1,4}^{(0)} = 0 \\ & u_{2i-1,2}^{(0)} - 4u_{2i-1,3}^{(0)} + 6u_{2i-1,3}^{(0)} + u_{2i-1,4}^{(0)} + \omega_{2i-1,3}^{(0)} + u_{2i-1,4}^{(0)} = 0 \\ & u_{2i-1,2}^{(0)} - 4u_{2i-1,3}^{(0)} + 6u_{2i-1,3}^{(0)} + 4u_{2i-1,3}^{(0)} + u_{2i-1,4}^{(0)} = 0 \\ & u_{2i-1,2n-3}^{(0)} + 6u_{2i-1,2n-2}^{(0)} - (2 + \omega h^3) u_{2i-1,2n-1}^{(0)} + u_{2i-1,2n-1}^{(0)} \\ & u_{2i-1,2n-3}^{(0)} + 6u_{2i-1,2n-2}^{(0)} - (2 + \omega h^3) u_{2i-1,2n-1}^{(0)} + u_{2i-1,2n-1}^{(0)} \\ & u_{2i-1,2n-1}^{(0)} + u_{2i-1,2n-1}^{(0)} \\ & (1 - (\omega/2) h^3) u_{2i-1,1}^{(0)} + u_{2i-1,2n-1}^{(0)} + u_{2i-1,2n-1}^{(0)} + u_{2i-1,2n-1}^{(0)} \\ & (1 - (\omega/2) h^3) u_{2i-1,1}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,3n-1}^{(0)} + u_{2i-1,3n-1}^{(0)} \\ & u_{2i-1,2n-1}^{(0)} \\ & (1 - (\omega/2) h^3) u_{2i-1,1}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,3n-1}^{(0)} + \frac{2(\phi_{2i-1,1}^{(0)} + \phi_{2i-1,3}^{(0)})}{u_{2i-1,2n-1}^{(0)}} \\ & (1 - (\omega/2) h^3) u_{2i-1,2n-2}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,3n-1}^{(0)} + \frac{2(\phi_{2i-1,1}^{(0)} + \phi_{2i-1,3}^{(0)})}{u_{2i-1,2n-1}^{(0)}} \\ & (1 - (\omega/2) h^3) u_{2i-1,2n-2}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,3n-1}^{(0)} + \frac{2(\phi_{2i-1,1}^{(0)} + \phi_{2i-1,3}^{(0)})}{u_{2i-1,2n-3}^{(0)} + u_{2i-1,2n-1}^{(0)} + \frac{2(\phi_{2i-1,1}^{(0)} + \phi_{2i-1,3}^{(0)})}{u_{2i-1,2n-2}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,2n-1}^{(0)} + \frac{2(\phi_{2i-1,1}^{(0)} + \phi_{2i-1,2n-1}^{(0)})}{u_{2i-1,2n-2}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,2n-1}^{(0)} + \frac{2(\phi_{2i-1,1}^{(0)} + \phi_{2i-1,2n-1}^{(0)})}{u_{2i-1,2n-3}^{(0)} + u_{2i-1,2n-2}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,2n-1}^{(0)} + \frac{2(\phi_{2i-1,1}^{(0)} + \phi_{2i-1,2n-1}^{(0)})}{u_{2i-1,2n-3}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,2n-3}^{(0)} + \frac{2(\phi_{2i-1,1}^{(0)} + \phi_{2i-1,2n-1}^{(0)})}{u_{2i-1,2n-2}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,2n-3}^{(0)} + \frac{2(\phi_{2i-1,2n-3}^{(0)} + \phi_{2i-1,2n-1}^{(0)})}{u_{2i-1,2n-3}^{(0)} + (1 - (\omega/2) h^3) u_{2i-1,2n-1}^{(0)} + \frac{2(\phi_{2i-1,2n-3}^{(0)}$$

convergence the components of the limit vector are certainly equal to the components of the solution of our system of difference equations corresponding to the crossed points.

The actual calculation proceeds in this way: First of all we calculate the boundary condition at the circled points. When the (k - 1)-st approximation  $\mathbf{v}^{(k-1)}$  is already known, we calculate the right-hand side of (18). Let us denote it by  $\chi^{(k-1)}$ . Then it is necessary to find  $\mathbf{v}^{(k)}$  in such a way that  $\mathbf{L}_{\omega}\mathbf{v}^{(k)} = \chi^{(k-1)}$ . Consequently, on every straight line of the first sort y = (2j - 1)h (j = 1, 2, ..., n), the system of equations (19) from the page 129 must be fulfilled, where

$$\varphi_{i,2j-1}^{(k)} = u_{i,2j-2}^{(k)} - \left(2 + \omega h^2\right) u_{i,2j-1}^{(k)} + u_{i,2j}^{(k)};$$

if (l, m) is an index corresponding to a crossed point then  $u_{lm}^{(k)}$  is the due component of the vector  $\mathbf{v}^{(k)}$  while in the opposite case  $u_{lm}^{(k)}$  is the value found by calculating so that to fulfil (16) or (17). Then the values  $u_{lm}^{(k)}$  must fulfil, on every straight line of the first sort x = (2i - 1)h (i = 1, 2, ..., n), the system (20) from the page 130. The system (19) is the difference approximation of the differential problem

$$\frac{\partial^2 \varphi^{(k)}(x, y)}{\partial x^2} - \omega \varphi^{(k)}(x, y) = \chi^{(k-1)}(x, y)$$
$$\varphi^{(k)}(0, y) = \frac{\partial^2 \psi(0, y)}{\partial y^2} - \omega \psi(0, y)$$
$$\varphi^{(k)}(1, y) = \frac{\partial^2 \psi(1, y)}{\partial y^2} - \omega \psi(1, y)$$

for y = (2j - 1) h (j = 1, 2, ..., n) and, similarly, (20) approximates the problem

$$\frac{\partial^2 u^{(k)}(x, y)}{\partial y^2} - \omega u^{(k)}(x, y) = \varphi^{(k)}(x, y)$$
$$u(x, 0) = \psi(x, 0) \quad u(x, 1) = \psi(x, 1)$$

for x = (2i - 1)h (i = 1, 2, ..., n).

We arrange the systems (19) and (20) for the calculation so that we substitute into the second and third and into the last but one and last but two equations respectively from the first and from the last equations. Then we substract from the third, fifth, ..., (2n - 1)-st equation always both the neighbouring equations and divide then by minus two. So the systems assume the form referred to on the page 129 and on the page 130. The systems (19') and (20') have both a diagonally dominant matrix and so the vector  $\mathbf{v}^{(k)}$  can be found unambiguously as their solution. Now we find the matrix expression of the operator  $L_{\omega}$ . To this aim, let us denote

$$\mathbf{v}_{2i-1}^{(k)} = \begin{bmatrix} u_{2i-1,1}^{(k)} \\ u_{2i-1,3}^{(k)} \\ \vdots \\ u_{2i-1,2n-1}^{(k)} \end{bmatrix}, \quad \mathbf{u}_{2i-1}^{(k)} = \begin{bmatrix} u_{2i-1,2}^{(k)} \\ u_{2i-1,4}^{(k)} \\ \vdots \\ u_{2i-1,2n-2}^{(k)} \end{bmatrix}, \quad \boldsymbol{\varphi}_{2i-1}^{(k)} = \begin{bmatrix} \varphi_{2i-1,1}^{(k)} \\ \varphi_{2i-1,3}^{(k)} \\ \vdots \\ \varphi_{2i-1,3}^{(k)} \\ \vdots \\ \varphi_{2i-1,3}^{(k)} \end{bmatrix}.$$

The indices of the components of the vectors  $\mathbf{v}_{2i-1}^{(k)}$ ,  $\varphi_{2i-1}^{(k)}$  correspond to the crossed points on the straight line of the first sort x = (2i - 1)h and similarly the vector  $\mathbf{u}_{2i-1}^{(k)}$  involves the circled points. In the following we shall try to eliminate the components of the vector  $\mathbf{u}_{2i-1}^{(k)}$  from the system (20'). Let us denote the rectangular matrix of n - 1 rows and n columns

$$\boldsymbol{J} = \begin{bmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & 1 & \\ & & \ddots & \ddots & \\ & & 1 & 1 \end{bmatrix}.$$

The odd equations of the system (20') give

(21) 
$$-(2+\omega h^2)\mathbf{v}_{2i-1}^{(k)} = -\mathbf{J}^{\mathsf{T}}\mathbf{u}_{2i-1}^{(k)} + \boldsymbol{\varphi}_{2i-1}^{(k)} + \mathbf{r}_i,$$

where  $r_i$  is a vector independent of k. Similarly, the even equations from (20') yield

(22) 
$$-2\boldsymbol{u}_{2i-1}^{(k)} = -\left(1 - \frac{\omega}{2}h^2\right)\boldsymbol{J}\boldsymbol{v}_{2i-1}^{(k)} + \frac{1}{2}\boldsymbol{J}\boldsymbol{\varphi}_{2i-1}^{(k)}.$$

When we substitute  $\boldsymbol{u}_{2i-1}^{(k)}$  from (22) into (21) we obtain

(23) 
$$\left\{-(2+\omega h^2)\mathbf{I}+\frac{1}{2}\left(1-\frac{\omega}{2}h^2\right)\mathbf{J}^{\mathsf{T}}\mathbf{J}\right\}\mathbf{v}_{2i-1}^{(k)}=\left(\mathbf{I}+\frac{1}{4}\mathbf{J}^{\mathsf{T}}\mathbf{J}\right)\varphi_{2i-1}^{(k)}+\mathbf{r}_i.$$

The matrix

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} 1 & 1 & & & \\ 1 & 2 & 1 & & \\ 1 & 2 & 1 & & \\ & 1 & 2 & 1 & \\ & & \ddots & & \\ & & & 1 & 2 & 1 \\ & & & & 1 & 1 \end{bmatrix}$$

.

is of the order *n*, positive semidefinite, of the rank n - 1 and thus  $I + \frac{1}{4}J^{T}J$  is regular

and we may multiply the system (23) by the inverse matrix. We obtain after an easy arrangement

$$(I + \frac{1}{4}J^{\mathsf{T}}J)^{-1} \{ -(2 + \omega h^2) (I + \frac{1}{4}J^{\mathsf{T}}J) + J^{\mathsf{T}}J \} \mathbf{v}_{2i-1}^{(k)} = \varphi_{2i-1}^{(k)} + (I + \frac{1}{4}J^{\mathsf{T}}J)^{-1} \mathbf{r}_i ,$$

that is

(24) 
$$\left\{-\left(2+\omega h^{2}\right)\mathbf{I}+\left(\mathbf{I}+\frac{1}{4}\mathbf{J}^{\mathsf{T}}\mathbf{J}\right)^{-1}\mathbf{J}^{\mathsf{T}}\mathbf{J}\right\}\mathbf{v}_{2i-1}^{(k)}=\boldsymbol{\varphi}_{2i-1}^{(k)}+\left(\mathbf{I}+\frac{1}{4}\mathbf{J}^{\mathsf{T}}\mathbf{J}\right)^{-1}\mathbf{r}_{i}$$

Let us denote the matrix on the left-hand side of (24) by  $\mathbf{A}_{\omega}$ . This matrix is a function of  $\mathbf{J}^{\mathsf{T}}\mathbf{J}$  and therefore if  $\lambda$  is an eigenvalue of  $\mathbf{J}^{\mathsf{T}}\mathbf{J}$  then  $\psi(\lambda) = -(2 + \omega h^2) + \lambda/(1 + \frac{1}{4}\lambda)$  is an eigenvalue of  $\mathbf{A}_{\omega}$ . We find easily that  $0 \leq \lambda < 4$ . The function  $\psi$  is increasing and therefore  $-(2 + \omega h^2) \leq \psi(\lambda) < -\omega h^2$ . Hence for positive  $\omega$ ,  $\mathbf{A}_{\omega}$  is negative definite.

Let us denote by  $\mathscr{A}_{\omega}$  the block diagonal matrix of *n* diagonal blocks



and let us denote

$$= \left[\varphi_{11}^{(k)}, \varphi_{13}^{(k)}, \dots, \varphi_{1,2n-1}^{(k)}, \varphi_{31}^{(k)}, \varphi_{33}^{(k)}, \dots, \varphi_{3,2n-1}^{(k)}, \dots, \varphi_{2n-1,1}^{(k)}, \varphi_{2n-1,3}^{(k)}, \dots, \varphi_{2n-1,2n-1}^{(k)}\right]^{\mathsf{T}}$$

 $\boldsymbol{\omega}^{(k)} =$ 

(indices corresponding to all crossed points). Then

$$\mathscr{A}_{\omega}\mathbf{v}^{(k)}=\boldsymbol{\varphi}^{(k)}+\boldsymbol{r}\,,$$

where  $\mathbf{r}$  is a vector independent of k.

In the same way we obtain from (19')

$$\mathbf{P}\mathscr{A}_{\omega}\mathbf{P}\boldsymbol{\varphi}^{(k)} = \boldsymbol{\chi}^{(k-1)} + \mathbf{s} ,$$

where **s** is a vector independent of k and **P** is the permutation matrix analogous to the matrix from the first section. The operator  $\mathbf{L}_{\omega}$  evidently does not depend on the exchange of axes x ans y, therefore  $\mathscr{A}_{\omega}\mathbf{P}\mathscr{A}_{\omega}\mathbf{P} = \mathbf{P}\mathscr{A}_{\omega}\mathbf{P}\mathscr{A}_{\omega}$  and it is according to our lemma a positive definite matrix for  $\omega \ge 0$ . Further, the proof of convergence  $\mathbf{v}^{(k)}$  is completely analogous as in the continuous case.

In the final part of the paper I should like to offer a very inaccurate consideration leading to an estimation of the optimal choice of  $\omega$ . When derivating the iterative formulas of the method we started from the symbolic formula (9), which can be obtained from the original equation (1) in the following way. The equation is rewritten in the form

$$-\omega\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = -\omega\Sigma u - \omega\chi$$

and the term

$$\omega^2 u + \frac{\partial^4 u}{\partial x^2 \partial v^2}$$

is added to both its sides.

When choosing  $\omega$  great enough, the member  $\partial^4 u/(\partial x^2 \partial y^2)$ , which makes the decomposition of the operator on the left-hand side possible, will be relatively small and we emphasize the original equation.

But if we choose  $\omega$  too great, we do not suppress only this expression but also the original equation and (9) is practically transformed to the identity. It seems to be optimal to choose  $\omega$  in the form C/h or  $C/h^2$ , where C is a positive constant.

However, we have to take into account the inner capacity of the computer in order to choose C only so great that the valid digits of  $\Sigma_{ij}$  from the expression  $\omega I - \Sigma$  in formula (8) or (18) be within the scope of the computer considered.

At present the author is preparing a series of numerical experiments which test the advantages of the methods, speed of convergence, sensibility towards truncation errors and the choice of optimal parameter. The results obtained correspond to the considerations presented in the paper and will be eventually the topics of another publication.

Finally, I should like to express my thanks to Prof. Ivo Marek DrSc. for valuable advice concerning the topic of the paper.

#### References

- [1] R. S. Varga: Matrix Iterative Analysis. Prentice Hall, Englewood Cliffs, New Jersey, 1962.
- [2] A. S. Householder: The Theory of Matrices in Numerical Analysis. Blaisdell Publishing Company, New York-Toronto-London, 1964.

### Souhrn

## O JEDNÉ ITERAČNÍ METODĚ ŘEŠENÍ DIRICHLETOVY ÚLOHY

### EMIL HUMHAL

Práce se zabývá iteračním řešením soustav lineárních algebraických rovnic vzniklých při diferenčním řešení eliptické rovnice

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - \Sigma(x, y) u = \chi(x, y)$$

na čtverci  $\langle 0, 1 \rangle \times \langle 0, 1 \rangle$  s okrajovou podmínkou  $u(x, y) = \psi(x, y)$  na hranici.

Studuje se konvergence dvou metod, které se dostanou vhodným diferenčním přepisem úlohy

$$\left(\frac{\partial^2}{\partial x^2} - \omega I\right) \left(\frac{\partial^2}{\partial y^2} - \omega I\right) u^{(k)} = \frac{\partial^4 u^{(k-1)}}{\partial x^2 \partial y^2} + \omega^2 u^{(k-1)} - \omega \Sigma u^{(k-1)} - \omega \chi$$

a okrajovou podmínkou  $u^{(k)}(x, y) = \psi(x, y)$  na hranici, pro k = 1, 2, ..., přičemž  $\omega$  je vhodný reálný parametr. Dokazuje se konvergence pro  $\omega \ge \max_{\substack{0 \le x \le 1 \\ 0 \le y \le 1}} \Sigma(x, y).$ 

Author's address: Emil Humhal, prom. mat., Fakulta jaderná a fysikálně inženýrská ČVUT, katedra matematiky, Husova 5, 110 00 Praha 1.