TW

stichting
mathematisch
centrum

MC

AFDELING TOEGEPASTE WISKUNDE TW 129/71 AUGUST

TW

P.J. VAN DER HOUWEN and T. ZWAAGSTRA-EERBEEK

ON THE NUMERICAL SOLUTION OF THE TRICOMI PROBLEM

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

Contents

	page
1. Introduction	1
2. The Tricomi problem	2
3. Numerical methods	3
4. The difference analogue of Fillippov	4
5. Iterative solution by the S.O.R. method	9
6. Successive over relaxation with y-dependent relaxation factor	22
7. Numerical solution of the Tricomi problem	28
8. References	31

1. Introduction

The purpose of the investigations presented in this report was to learn about the properties of relaxation methods when applied to a non-elliptic boundary value problem. Such a problem is the Tricomi problem which is partly elliptic, parabolic and hyperbolic. We took the Tricomi problem with a rectangular boundary curve in the elliptic region as a model problem.

Our starting point was the iterative method given by Fillippov [3]. This method is, in fact, Gauss-Seidel's method applied to a vector equation which represents a very special discrete analogue of the analytical problem. We took the same discretization, but applied the successive overrelaxation method of Young [14]. However, the matrix A associated to the vector equation of Fillippov does not have "property (A)" so that Young's theory could not be applied. Therefore, we first considered Tricomi's equation in that region where it is elliptic, i.e. we considered a Dirichlet problem. The matrix A restricted to the elliptic region does have "property (A)" and by applying the theory of Young we found a formula for the optimal relaxation factor as a function of the grid distance h. We also applied the technique of Garabedian for estimating relaxation factors; this resulted in a formula not only depending on h but also on the coordinates of the grid points. Our experiments showed, however, that both formulae give a comparable rate of convergence.

We then returned to the complete Tricomi problem. We did experiments in which the relaxation factor used in the elliptic region was different from the one used in the hyperbolic region. The pair of values which appeared to yield the largest rate of convergence turned out to be largely different. The elliptic one is close to the optimal value holding for the Dirichlet problem and tends to 2 as $h \to 0$, while the hyperbolic one drops below 1 as $h \to 0$. Thus, we have overrelaxation in the elliptic region and underrelaxation in hyperbolic region.

Finally, we did experiments with a fixed relaxation factor for both the elliptic and hyperbolic region. When this factor equals 1, the method reduces to Gauss-Seidel's method and is identical to Fillippov's original method. We found a considerable lower rate of convergence (depending on h). By increasing the relaxation factor, somewhere between the optimal values for the elliptic region and hyperbolic region, we got a more rapid convergence, but still below the rate of convergence of the "over-underrelaxation" method.

2. The Tricomi problem

The Tricomi problem is a boundary value problem for the equation

$$y\ \psi_{xx} + \psi_{yy} = f\ , \qquad (2.1)$$

where the domain R, in which the equation is to be solved, consists of an elliptic region R^+ in the upper or elliptic halfplane and a hyperbolic region R^- in the lower or hyperbolic halfplane. R^- is bounded by two characteristics of the equation (see figure 2.1).

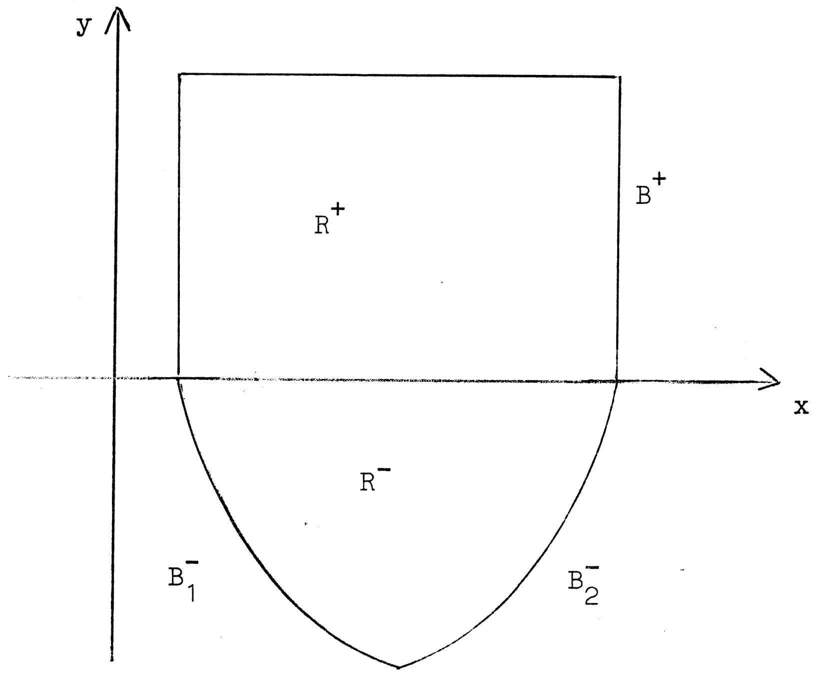

fig. 2.1. The Tricomi boundary value problem.

Equation (2.1) is said to be of mixed type, as it is elliptic for $y > 0$, parabolic for $y = 0$ and hyperbolic for $y < 0$.

It can be proved that (2.1) has a unique solution when boundary values are prescribed at the complete elliptic part B^+ of the boundary and at one of the characteristics, say B_1^-, in the hyperbolic plane (see for instance reference [6]).

In this paper, we shall consider, as a model problem, the case where the elliptic region R^+ is bounded by three sides of a square of side 1 (see figure 2.2). The characteristics B_1^- and B_2^- are then given by

$$(2.2)\qquad \begin{aligned} B_1^- &: \quad x - \tfrac{2}{3}(-y)^{3/2} = 0 \ , \\ B_2^- &: \quad x + \tfrac{2}{3}(-y)^{3/2} = 1 \ . \end{aligned}$$

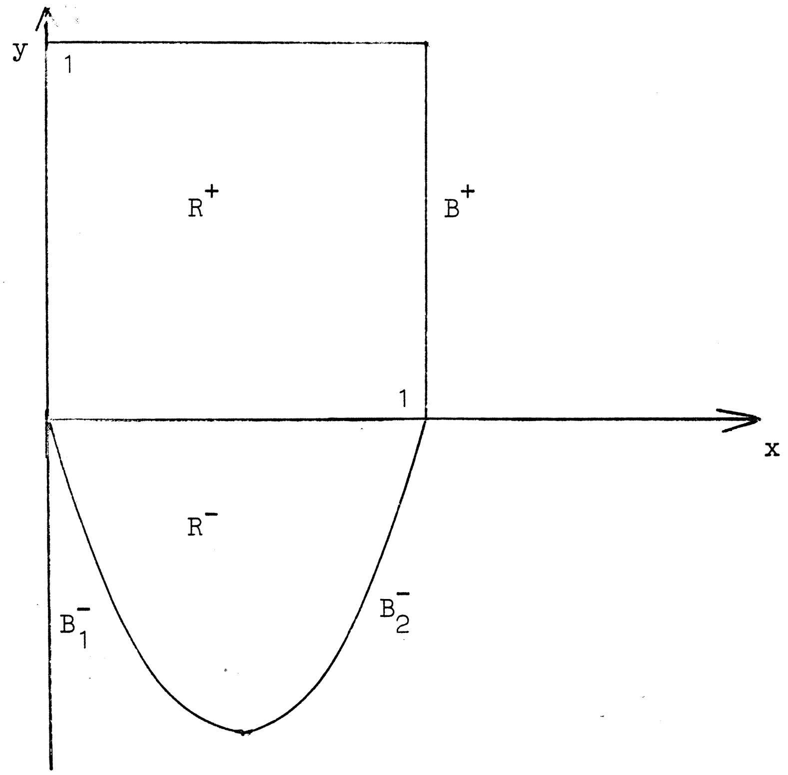

fig. 2.2. The model problem.

3. Numerical methods

In reference [6], where analytical aspects of the Tricomi problem are considered, it was pointed out that solutions of equation (2.1) can only be obtained in an approximate manner. There are two important (numerical) approaches to construct solutions; expansion in a series of particular solutions and difference methods.

The first method of solution can be found in Bergman [1], Guderly and Yoshihara [5], and in Ovsiannikov [11]. In some cases this leads to an effective solution of the Tricomi problem.

Difference methods, however, have turned out to be more successful. We mention the numerical calculations of Vincenti and Wagoner [13] and the theoretical justification of their scheme by Mussman. Further the work of Levy [9], who gave the first rigorous treatment by difference methods of the Tricomi problem. In the papers just mentioned, the boundary condition at the characteristic B_1^- was transformed in a boundary condition at the parabolic line $y = 0$. Then the problem is an elliptic boundary value problem with a complicated boundary condition at the parabolic part of the boundary.

An approach which does not use the transformation of the hyperbolic boundary condition was given by Chu [2]. However, his method only applies to rectangular domains R. Therefore, a complicated transformation is necessary to map R on a rectangle.
Finally, a method which applies to any region which is, in the hyperbolic region, bounded by two characteristics, was given by Fillippov [3]. It is this method which will be considered in some detail in the following section. In the subsequent sections, procedures will be investigated in order to accelerate the convergence of Fillippov's method.

4. The difference analogue of Fillippov

Fillippov used the following grid in the (x,y)-plane (see figure 4.1): the mesh points in the hyperbolic plane are defined as the intersections of the characteristics originating from equally spaced points at the parabolic line;
the elliptic mesh points are obtained by reflecting the hyperbolic points with respect to the parabolic line and by completing these elliptic points to obtain a rectangular grid in the elliptic plane.

This grid is the starting point of the discretization method of Fillippov. The next step is to replace the derivatives $\partial\psi^2/\partial x^2$ and $\partial\psi^2/\partial y^2$ by difference quotients defined at the mesh points. For the sake of simplicity we shall assume that the boundary B^+ runs through the mesh points just defined.

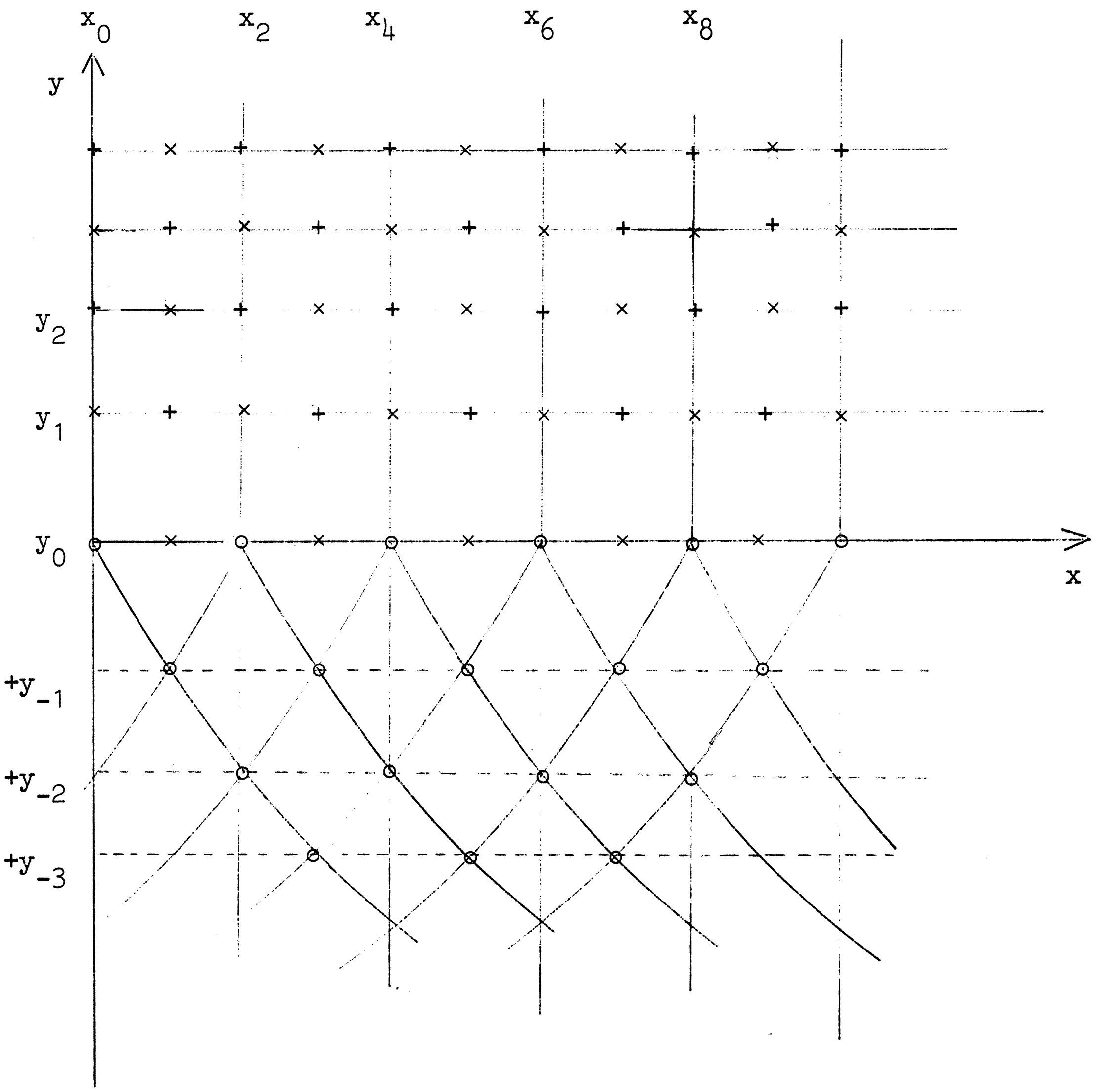

fig. 4.1.

o hyperbolic mesh points

+ elliptic mesh points obtained by reflection of the hyperbolic points

× elliptic mesh points obtained by completing the elliptic points

The elliptic difference formula.

Fillippov used a five-point formula which expresses the value of $\psi(x,y)$ in an elliptic point (see figure 4.1) in the values of $\psi(x,y)$ at the four neighbouring points x, and vice versa.
Let us define the grid parameters (see figure 4.1)

$$(4.1)\qquad \begin{aligned} h &= x_n - x_{n-1} \ , \\ l_m &= y_m - y_{m-1} \ . \end{aligned}$$

Then, it is easily verified that as $h \to 0$ the formula

$$(4.2)\qquad \begin{aligned} &\frac{y_m}{h^2}\,(\psi(n+1,m) + \psi(n-1,m)) + \frac{2}{l_m + l_{m+1}}\,(\frac{1}{l_{m+1}}\,\psi(n,m+1) + \frac{1}{l_m}\,\psi(n,m-1)) \\ &- 2(\frac{y_m}{h^2} + \frac{1}{l_m l_{m+1}})\ \psi(n,m) = f(n,m) \ , \end{aligned}$$

$\psi(n,m)$ and $f(n,m)$ being the values of the functions ψ and f at the mesh point (n,m), is a consistent approximation to equation (2.1) at an internal elliptic point.

The parabolic difference formula.

At a parabolic mesh point we simply have the relation (see figure 4.1)

$$(4.3)\qquad \frac{1}{y_j^2}\,(\psi(n,j) - 2\psi(n,0) + \psi(n,-j)) = f(n,0) \ ,$$

where $j = 1$ when n is odd and $j = 2$ when n is even.

The hyperbolic difference formula.

In the hyperbolic region Fillippov used four mesh points to approximate the differential equation (see figure 4.2).

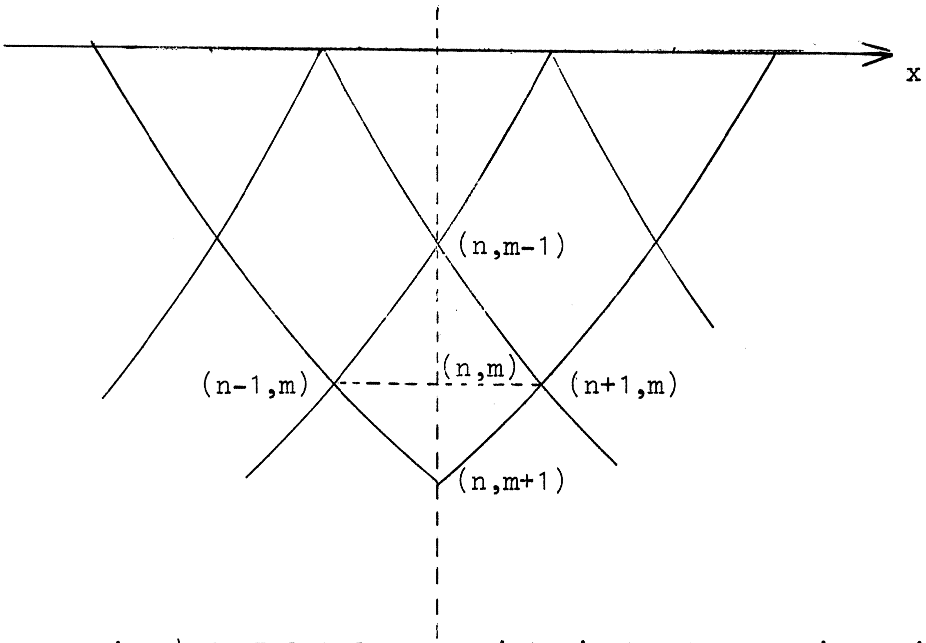

fig. 4.2. Related mesh points in the hyperbolic region.

It can be proved that as $h \to 0$ the formula

$$(4.4)\qquad -\frac{1}{l_m l_{m+1}}(\psi(n+1,m)+\psi(n-1,m))+\frac{2}{l_m+l_{m+1}}\left(\frac{1}{l_m}\psi(n,m+1)+\frac{1}{l_{m+1}}\psi(n,m-1)\right) = f(n,m)$$

approximates the differential equation at the hyperbolic point (n,m). This formula is not so easily verified as formulae (4.2) and (4.3). Therefore, some explanation will be given. Let us expand the left hand side of (4.4) in a Taylor series with respect to the hyperbolic point (n,m). We then get

$$(4.5)\qquad \frac{-h^2}{l_m l_{m+1}}\psi_{xx}(n,m)+\psi_{yy}(n,m) = f(n,m)+s(h) ,$$

where $s(h)$ tends to zero as $h \to 0$.
In order to prove that (4.5) converges to (2.1) as $h \to 0$, we have to show that

$$(4.6) \qquad \lim_{h \to 0} (h^2 - l_m l_{m+1} y_m) = 0.$$

For that purpose we express h in terms of y_m and m.
From (2.2) we derive that the two characteristics passing through the point $(x_n, -y_m)$ are given by

$$x \pm \frac{2}{3} (-y)^{3/2} = x_n \pm \frac{2}{3} y_m^{3/2} \ .$$

The values of the coordinates of the points where these characteristics intersect the parabolic line differ by 2mh. Hence

$$(4.7) \qquad 2mh = \frac{4}{3} y_m^{3/2} \ .$$

From (4.7) and the mean value theorem we find for small values of h

$$h = \frac{2}{3} (y_{m+1}^{3/2} - y_m^{3/2}) \quad \frac{2}{3} l_{m+1} \frac{3}{2} y_m^{1/2} \ .$$

This proves relation (4.6).

We now define a vector u whose components have, in some order, the values which the function $\psi(x,y)$ assumes at the net points (including the boundary points). Furthermore, we define a vector f composed of the values of $\psi(x,y)$ at the boundary points of the boundary part $B^+ + B_1^-$, and the values of f(x,y) at the remaining mesh points. The components of f are arranged in the same order as the components of u.
With these definitions the discrete analogue of the Tricomi boundary value problem can be written as a vector equation

$$(4.8) \qquad Au = f \ .$$

At the internal net points the matrix A is defined by the difference formulae (4.2) - (4.4). At the boundary points where u is prescribed, A is the identity operator.

5. Iterative solution by the SOR method

In this section we give the results obtained by applying the SOR or successive overrelation method (Frankel [4], Young [14]) to equation (4.8). The definition of this method is most easily given by writing the matrix A in the form

$$A = C - E - F , \tag{5.1}$$

where C is a diagonal matrix, whose entries are the diagonal elements of A, and E and F are respectively strictly lower and upper triangular matrices, whose entries are the negatives of the entries of A respectively below and above the main diagonal of A. We now define the SOR method by the recurrence relation

$$u_{k+1} = (1-\Omega)u_k + \Omega C^{-1}(Eu_{k+1}+Fu_k+f) , \tag{5.2}$$

where u_0 is an arbitrary initial approximation and Ω is the relaxation factor with values between 0 and 2.

When $\Omega = 1$ the SOR method reduces to Gauss-Seidel's method. Fillippov proved the convergence of this method when applied to the Tricomi boundary value problem. However, the rate of convergence is small and we have tried to accelerate the convergence by choosing more appropriate values for Ω.

If the matrix A should possess what Young [14] called "property(A)" one can give relations for the optimal value of Ω. However, the matrix A as defined in the preceding section does not possess property(A), irrespective of the order of the components u. The difficulties arise in the hyperbolic region where the coupling of the components of u is strong. Moreover, it may be remarked that in cases where A does have this property, the optimal value of Ω is related to the spectral radius of the matrix

$$B = I - C^{-1}A . \tag{5.3}$$

Hence, the problem is replaced by a not so simple eigenvalue problem. Furthermore, the complete different character of the matrix operator A

in the elliptic and hyperbolic region, respectively, suggests to use different values of Ω in R^+ and R^-, say Ω^+ and Ω^-. The optimal values of $\Omega^\pm$ have been obtained experimentally. First, we considered the problem in which the hyperbolic difference formulae are omitted and boundary values are prescribed at the parabolic line. Then, the optimal Ω^+ found for this Dirichlet problem were used in determining Ω^- for the hyperbolic region. At the parabolic line we used formula (4.3).
For the boundary values we took the values which the analytical solution

(5.4) $$\phi(x.y) = y^3 - 3x^2$$

assumes at B^+ and B_1^-.
The initial vector u_0 was set equal to 0.

<u>The Dirichlet problem</u>

We applied the SOR method to a square of side 1 (see figure 5.1),

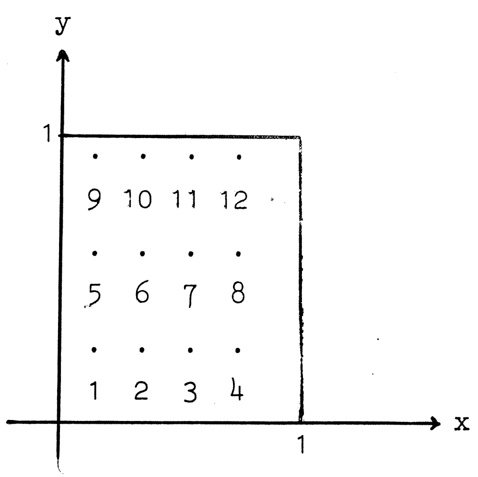

fig. 5.1 Dirichlet problem for $y\phi_{xx} + \phi_{yy} = 0$

where the boundary values are zero at the parabolic line and according to (5.4) at the remaining sides.
The order of the components of u were arranged as indicated in figure 5.1.
In the elliptic region the matrix A is given for h = 1/6 in fig. 5.2.

fig. 5.2. Matrix A_{ell} for h = 1/6.

A_{ell} =

1	-.285	0	0	0	-.271	0	0	0	0	0	0	0	0	0
-.285	1	-.285	0	0	0	-.271	0	0	0	0	0	0	0	0
0	-.285	1	-.285	0	0	0	-.271	0	0	0	0	0	0	0
0	0	-.285	1	-.285	0	0	0	-.271	0	0	0	0	0	0
0	0	0	-.285	1	0	0	0	0	-.271	0	0	0	0	0
-.224	0	0	0	0	1	-.254	0	0	0	-.267	0	0	0	0
0	-.224	0	0	0	-.254	1	-.254	0	0	0	-.267	0	0	0
0	0	-.224	0	0	0	-.254	1	-.254	0	0	0	-.267	0	0
0	0	0	-.224	0	0	0	-.254	1	-.254	0	0	0	-.267	0
0	0	0	0	-.224	0	0	0	-.254	1	0	0	0	0	-.267
0	0	0	0	0	-.234	0	0	0	0	1	-.252	0	0	0
0	0	0	0	0	0	-.234	0	0	0	-.252	1	-.252	0	0
0	0	0	0	0	0	0	-.234	0	0	0	-.252	1	-.252	0
0	0	0	0	0	0	0	0	-.234	0	0	0	-.252	1	-.252
0	0	0	0	0	0	0	0	0	-.234	0	0	0	-.252	1

Ω^+ \ k	10	20	30	40	50	60
0	0	0	0	0	0	0
0.1	0.063	0.056	0.051	0.048	0.045	0.043
0.2	0.12	0.099	0.088	0.080	0.074	0.069
0.3	0.17	0.14	0.12	0.11	0.10	0.094
0.4	0.21	0.17	0.15	0.13	0.13	0.12
0.5	0.26	0.20	0.18	0.17	0.16	0.15
0.6	0.30	0.24	0.21	0.20	0.20	0.19
0.7	0.34	0.28	0.26	0.25	0.24	0.24
0.8	0.39	0.33	0.31	0.30	0.29	0.29
0.9	0.44	0.39	0.37	0.36	0.36	0.36
1.0	0.50	0.47	0.45	0.45	0.45	0.44
1.1	0.59	0.58	0.58	0.57	0.56	0.47
1.2	0.75	0.78	0.79	0.68	0.54	0.45
1.3	1.00	1.10	0.93	0.70	0.56	0.46
1.4	0.78	0.83	0.85	0.68	0.57	0.47
1.5	0.61	0.63	0.65	0.65	0.55	0.46
1.6	0.45	0.47	0.47	0.47	0.49	0.46
1.7	0.32	0.33	0.33	0.33	0.34	0.34
1.8	0.20	0.20	0.21	0.20	0.21	0.21
1.9	0.093	0.092	0.094	0.091	0.096	0.099
2.0	-0.006	-0.008	-0.007	-0.010	-0.008	-0.006

Table 5.1. Rates of convergence for the Dirichlet problem with 1/6.

Further, as an estimate of the rate of convergence after k iterations we used the value of

$$(5.5) \qquad R^{*}(k) = -\frac{1}{k} \ln \left[\frac{||Au_k - f||}{||Au_0 - f||} \right] ,$$

where $||\ \ ||$ denotes the Euclidean norm in the space of iterates u_k (see [7], p. 7). In table 5.1, 5.2 and 5.3 the values of $R^{*}(k)$ are given for some values of h and Ω^+.

Table 5.2. Rates of convergence for the Dirichlet problem with h = 1/12

Ω^+ \ k	10	20	30	40	50	60	70	80	90	100	110	120
0	0	0	0	0	0	0	0	0	0	0	0	0
0.1	0.053	0.045	0.040	0.035	0.032	0.029	0.027	0.026	0.024	0.023	0.022	0.021
0.2	0.094	0.073	0.060	0.052	0.047	0.043	0.037	0.037	0.035	0.033	0.032	0.031
0.3	0.13	0.093	0.076	0.066	0.059	0.054	0.050	0.047	0.044	0.042	0.040	0.039
0.4	0.15	0.11	0.090	0.078	0.070	0.064	0.060	0.056	0.053	0.050	0.048	0.046
0.5	0.18	0.13	0.10	0.090	0.081	0.074	0.069	0.065	0.061	0.058	0.056	0.054
0.6	0.20	0.14	0.12	0.10	0.092	0.084	0.078	0.074	0.070	0.067	0.065	0.063
0.7	0.22	0.16	0.13	0.11	0.10	0.094	0.088	0.084	0.081	0.078	0.076	0.074
0.8	0.25	0.18	0.14	0.13	0.11	0.11	0.10	0.096	0.093	0.090	0.088	0.087
0.9	0.27	0.19	0.16	0.14	0.13	0.12	0.11	0.11	0.11	0.11	0.10	0.10
1.0	0.29	0.21	0.17	0.16	0.15	0.14	0.13	0.13	0.13	0.12	0.12	0.12
1.1	0.32	0.23	0.19	0.18	0.17	0.16	0.15	0.15	0.15	0.15	0.14	0.14
1.2	0.34	0.25	0.22	0.20	0.19	0.19	0.18	0.18	0.18	0.18	0.17	0.17
1.3	0.37	0.28	0.25	0.24	0.23	0.23	0.23	0.22	0.22	0.22	0.22	0.21
1.4	0.38	0.32	0.31	0.30	0.30	0.29	0.29	0.29	0.29	0.26	0.24	0.22
1.5	0.35	0.42	0.42	0.43	0.43	0.43	0.38	0.33	0.29	0.26	0.24	0.22
1.6	0.27	0.41	0.43	0.44	0.46	0.44	0.37	0.33	0.29	0.26	0.24	0.22
1.7	0.18	0.29	0.30	0.31	0.33	0.32	0.33	0.32	0.29	0.26	0.24	0.21
1.8	0.095	0.19	0.18	0.19	0.20	0.20	0.21	0.21	0.21	0.21	0.21	0.21
1.9	0.017	0.09	0.077	0.087	0.093	0.091	0.096	0.095	0.097	0.098	0.097	0.098
2.0	divergent											

Table 5.3. Rates of convergence for the Dirichlet problem with h = 1/18.

Ω^+ \ k	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.1	.050	.043	.038	.034	.030	.028	.025	.024	.022	.021	.020	.019	.018	.017	.016	.016	.015	.015
0.2	.090	.069	.056	.048	.042	.038	.035	.032	.030	.028	.026	.025	.024	.023	.022	.021	.020	.020
0.3	.12	.087	.069	.058	.051	.046	.042	.038	.036	.034	.032	.030	.029	.028	.027	.026	.025	.024
0.4	.15	.10	.079	.067	.058	.052	.048	.044	.041	.039	.037	.035	.034	.033	.031	.030	.029	.028
0.5	.17	.11	.089	.074	.065	.059	.054	.050	.047	.044	.042	.040	.038	.037	.036	.034	.033	.032
0.6	.19	.12	.097	.082	.072	.065	.060	.056	.052	.050	.047	.045	.043	.041	.040	.039	.037	.036
0.7	.20	.14	.11	.090	.080	.072	.066	.062	.058	.055	.052	.050	.048	.046	.045	.043	.042	.041
0.8	.22	.15	.12	.099	.087	.079	.073	.068	.064	.060	.057	.055	.053	.047	.049	.048	.047	.046
0.9	.24	.16	.13	.11	.095	.086	.079	.074	.070	.067	.064	.061	.060	.058	.056	.055	.054	.053
1.0	.26	.17	.14	.12	.10	.094	.087	.082	.077	.074	.071	.069	.067	.066	.064	.063	.062	.061
1.1	.27	.18	.15	.13	.11	.10	.096	.091	.087	.083	.081	.079	.077	.075	.074	.073	.072	.071
1.2	.29	.20	.16	.14	.12	.11	.11	.10	.098	.095	.093	.091	.089	.088	.086	.085	.084	.083
1.3	.31	.21	.17	.15	.14	.13	.12	.11	.11	.11	.11	.11	.10	.10	.10	.10	.10	.10
1.4	.33	.23	.19	.17	.16	.15	.14	.14	.14	.13	.13	.13	.13	.13	.13	.12	.12	.12
1.5	.33	.25	.21	.20	.19	.18	.18	.17	.17	.17	.17	.17	.16	.16	.16	.16	.15	.15
1.6	.28	.25	.26	.25	.25	.24	.24	.24	.24	.24	.23	.22	.20	.19	.17	.16	.15	.14
1.7	.20	.18	.27	.28	.30	.31	.31	.31	.28	.26	.23	.21	.20	.18	.17	.16	.15	.14
1.8	.11	.10	.19	.18	.18	.19	.19	.20	.20	.20	.21	.20	.19	.18	.17	.16	.15	.14
1.9	.027	.034	.092	.080	.082	.088	.088	.094	.092	.092	.096	.096	.097	.097	.099	.098	.098	.099
2.0	divergent																	

From these tables we may draw three important conclusions:

1. The interval I_k of optimal Ω^+ values increases with the number of iterations k. The intervals I_k are indicated in the tables.
2. The rates of convergence corresponding to $\Omega^+ \in I_k$ are slowly varying compared with the rates of convergence obtained for non-optimal values of Ω^+.
3. After a number of iterations, when the iteration process becomes stationary, the interval I_{k_1} contains a preceding interval I_{k_2}, i.e. if $k_2 < k_1$ then $I_{k_1} \supset I_{k_2}$.

From these conclusions it follows that one should use those values of Ω^+ which are in the interval I_k obtained as soon as the process becomes stationary. Thus, we may use

$$(5.6) \qquad \Omega^+(1/6) = 1.3, \quad \Omega^+(1/12) = 1.55, \quad \Omega^+(1/18) = 1.7.$$

There remains the problem how we can predict from these results the optimal Ω^+ values for other values of h. In the case of the Dirichlet problem for Laplace's equation, the theory of Young yields the relation

$$(5.7) \qquad \Omega \sim 2 - ah, \text{ as } h \to 0,$$

where a is a constant only depending on the region R and not on h. Hence, when a is experimentally determined for one (sufficiently small) value of h one can predict by (5.7) the optimal value of Ω for other values of h. It may be interesting to give the following heuristic application of the theory of Young to our Dirichlet problem. As is already remarked, the optimal value of Ω for matrices A having property (A) is related to the spectral radius $\sigma(B)$ of the matrix B. We have, infact,

$$(5.8) \qquad \Omega = 1 + \left[\frac{\sigma(B)}{1 + (1-\sigma^2(B))^{1/2}}\right]^2 .$$

Now, the matrix A corresponding to our Dirichlet problem does have property (A). Hence, there only remains the problem to determine $\sigma(B)$. It is well-known that by applying Gerschgorin's theorem (see e.g. Varga [12], p. 16), one can obtain an upperbound for $\sigma(B)$. In doing so, however, we obtain the non-interesting result $\sigma(B) \leq 1$ or $\Omega \leq 2$. In order to get an upperbound for $\sigma(B)$ less than 1, we proceed as follows. The matrix C^{-1} is a diagonal matrix with diagonal entries

$$-\frac{1}{2}\left(\frac{y_m}{h^2} + \frac{1}{l_m l_{m+1}}\right)^{-1}.$$

For $h \to 0$ we have (compare section 4)

$$l_m \sim h/\sqrt{y_m} ,$$

hence the diagonal entries of C^{-1} behave as

$$\frac{-1}{4y_m} h^2 .$$

From (5.3) it then follows that B has the form

$$B \sim I + O(h^2)A ,$$

$$\text{as } h \to 0$$

Therefore, it is expected that

$$\sigma(B) = 1 - a\, h^2 , \qquad (5.9)$$

$$\text{as } h \to 0 ,$$

where a is a positive constant.

On the other hand we derive from (5.8) the relation

$$\sigma(B) = \frac{2}{\Omega} \sqrt{\Omega - 1} . \qquad (5.8')$$

In order to estimate a by means of (5.8') and (5.9) we have done experiments with a finer grid of Ω^+ values. The results are given in table 5.1' and 5.3'. With the aid of (5.8') we deduce from these more detailed results that

$$.79 \le \sigma(B) \le .89 \qquad \text{as } h = \frac{1}{6} \,,$$

(5.10)

$$.975 \le \sigma(B) \le .980 \qquad \text{as } h = \frac{1}{18} \,,$$

Hence, by applying (5.9) we find for h = 1/6 and h = 1/18 respectively

(5.11) $$4.7 \le a \le 7.5, \quad 6.5 \le a \le 8.1 \,.$$

In our subsequent calculations we assume that $a \sim 7.29$ so that $\sigma(B)$ is given by the relation

(5.12) $$\sigma(B) = 1 - 7.29\, h^2 \,.$$

As an application we calculate $\Omega^+(1/12)$ and $\Omega^+(1/24)$ by means of (5.8) and (5.12) we find the values

$$\Omega^+(1/12) = 1.52 \quad , \quad \Omega^+(1/24) = 1.72 \,.$$

Tables 5.2 and 5.4 show that these predicted values are in agreement with numerical experiments.

This section is concluded with a survey of the rates of convergence found for the original Fillippov scheme, i.e. $\Omega^+ = 1$, and the S.O.R. method with an optimal relaxation factor (see table 5.5). This table clearly shows the superiority of the S.O.R. method.

Table 5.1' Rates of convergence for the Dirichlet problem with h = 1/16

Ω^+ \ k	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	200	210	220	230	240	250
1.20	.75	.78	.79	.68	.54	.45	.39	.34	.30	.27	.25	.23	.21	.19	.18	.17	.16	.15	.14	.14	.13	.12	.12	.11	.11
1.22	.80	.85	.87	.70	.56	.47	.40	.35	.31	.28	.26	.23	.22	.20	.19	.18	.17	.16	.15	.14	.13	.13	.12	.12	.11
1.24	.87	.96	.91	.68	.55	.46	.39	.34	.30	.27	.25	.23	.21	.20	.18	.17	.16	.15	.14	.14	.13	.12	.12	.11	.11
1.26	.98	1.20	.92	.69	.55	.46	.39	.35	.31	.28	.25	.23	.21	.20	.18	.17	.16	.15	.15	.14	.13	.13	.12	.12	.11
1.28	1.10	1.10	.92	.69	.55	.46	.40	.35	.31	.28	.25	.23	.21	.20	.18	.17	.16	.15	.15	.14	.13	.13	.12	.12	.11
1.30	1.00	1.10	.93	.70	.56	.46	.40	.35	.31	.28	.25	.23	.21	.20	.19	.17	.16	.15	.15	.14	.13	.13	.12	.12	.11
1.32	.98	1.00	.92	.69	.55	.46	.39	.34	.31	.28	.25	.23	.21	.20	.18	.17	.16	.15	.14	.14	.13	.13	.12	.11	.11
1.34	.94	.96	.91	.67	.55	.45	.39	.34	.30	.27	.25	.22	.21	.19	.18	.17	.16	.15	.14	.13	.13	.12	.12	.11	.11
1.36	.88	.91	.89	.68	.54	.45	.39	.34	.30	.27	.25	.23	.21	.19	.18	.18	.16	.15	.14	.14	.13	.12	.12	.11	.11
1.38	.82	.87	.89	.67	.54	.45	.39	.34	.30	.27	.25	.22	.21	.19	.18	.17	.16	.15	.14	.13	.13	.12	.12	.11	.11
1.40	.78	.83	.85	.68	.57	.47	.41	.35	.32	.28	.26	.24	.22	.20	.19	.18	.17	.16	.15	.14	.14	.13	.12	.12	.11

Table 5.3'. Rates of convergence for the Dirichlet problem with h = 1/18

k \ Ω^+	1.57	1.58	1.59	1.60	1.61	1.62	1.63	1.64	1.65	1.66	1.67	1.68	1.69	1.70	1.71	1.72	1.73	1.74	1.75
10	.30	.29	.29	.28	.27	.26	.25	.25	.24	.23	.22	.21	.20	.20	.19	.18	.17	.16	.15
20	.26	.26	.25	.25	.25	.24	.23	.23	.22	.21	.20	.20	.19	.18	.17	.16	.16	.15	.14
30	.24	.25	.25	.26	.27	.28	.29	.30	.31	.31	.30	.29	.28	.27	.26	.26	.25	.24	.23
40	.23	.24	.24	.25	.26	.27	.29	.30	.33	.33	.31	.30	.29	.28	.27	.26	.25	.24	.23
50	.22	.23	.24	.25	.26	.27	.28	.30	.33	.34	.33	.32	.31	.30	.28	.27	.26	.25	.24
60	.22	.23	.24	.25	.26	.27	.28	.31	.34	.35	.34	.33	.32	.31	.29	.28	.27	.26	.25
70	.22	.22	.23	.24	.25	.27	.28	.31	.24	.36	.35	.34	.32	.31	.30	.29	.27	.26	.25
80	.21	.22	.23	.24	.25	.27	.28	.31	.32	.32	.32	.32	.32	.31	.31	.29	.29	.27	.26
90	.21	.22	.23	.24	.25	.27	.28	.28	.29	.29	.29	.28	.29	.28	.28	.28	.28	.27	.26
100	.21	.22	.23	.24	.25	.26	.26	.26	.26	.26	.26	.26	.26	.26	.26	.26	.25	.25	.25
110	.21	.22	.23	.23	.24	.23	.24	.23	.23	.23	.23	.23	.23	.23	.23	.23	.23	.23	.23
120	.21	.21	.21	.22	.22	.21	.21	.21	.22	.21	.21	.21	.21	.21	.21	.21	.21	.21	.21

Table 5.4. Rates of convergence for the Dirichlet problem with h = 1/24.

Ω^+ \ k	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180
1.10	.26	.17	.13	.11	.096	.086	.079	.073	.068	.065	.062	.059	.057	.055	.053	.052	.051	.050
1.20	.28	.18	.14	.12	.10	.093	.085	.079	.075	.071	.068	.065	.063	.062	.060	.059	.057	.056
1.30	.29	.19	.15	.13	.11	.10	.093	.087	.083	.079	.076	.074	.072	.070	.069	.067	.066	.065
1.40	.31	.20	.16	.14	.12	.11	.10	.098	.094	.091	.088	.085	.084	.082	.080	.079	.078	.077
1.50	.32	.22	.18	.15	.14	.13	.12	.11	.11	.11	.10	.10	.10	.099	.098	.097	.096	.095
1.60	.28	.23	.19	.17	.16	.15	.14	.14	.14	.14	.13	.13	.13	.13	.13	.13	.13	.13
1.70	.21	.19	.18	.22	.21	.21	.21	.21	.20	.20	.20	.20	.19	.18	.17	.16	.15	.14
1.72	.19	.17	.17	.23	.24	.24	.24	.24	.24	.24	.23	.21	.20	.18	.17	.16	.15	.14
1.74	.17	.16	.15	.22	.24	.25	.25	.26	.25	.25	.23	.21	.20	.18	.17	.16	.15	.14
1.76	.16	.14	.14	.21	.22	.23	.23	.23	.23	.25	.23	.21	.20	.18	.17	.16	.15	.14
1.78	.14	.13	.12	.19	.20	.21	.21	.21	.21	.23	.22	.21	.19	.18	.17	.16	.15	.14
1.80	.12	.11	.11	.18	.18	.19	.19	.19	.19	.20	.20	.20	.19	.18	.17	.16	.15	.14
1.90	.036	.040	.041	.091	.086	.083	.086	.089	.089	.097	.094	.093	.091	.098	.096	.096	.096	.096

Table 5.5. Rates of convergence for $\Omega^{+} = 1$ and $\Omega^{+} = \Omega^{+}_{opt}$.

h	k	Fillippov	S.O.R.	gainfactor
1/6	30	.45	.93	2.07
	60	.44	.46	1.05
1/12	30	.17	.43	2.53
	60	.14	.44	3.14
	90	.13	.29	2.23
	120	.13	.22	1.69
1/18	30	.14	.31	2.21
	60	.09	.35	3.89
	90	.08	.29	3.62
	120	.07	.21	3.0

6. Successive overrelaxation with a y-dependent relaxation factor.

In the preceding section a fixed value of Ω was used in the elliptic region. We now show that applying Garabedian's technique leads us to relaxation factors which depend on the value of y_m of the grid point in which the iteration process is applied.

First, we introduce the vectors

(6.1) $$v_k = u_k - u \quad , \quad k = 0,1,2,\ldots \; ,$$

where u is the solution of Au = f. These error vectors satisfy the homogeneous scheme (compare (5.2))

(6.2) $$(1-\Omega C^{-1}E)\; v_{k+1} = (1-\Omega+\Omega C^{-1}F)\; v_k \; .$$

The vectors v_k may be interpreted as values of a vector function V(t) in the points $t = t_k = k\tau$, where τ is a time step of arbitrary length. When the successive vectors v_k are changing slowly we may write

$$v_{k+1} = V(t_k+\tau) \cong V(t_k) + \tau\dot{V}(t_k) \quad ,$$

and equation (6.2) transforms into the first order ordinary differential equation

(6.3) $$\tau(\Omega E-C)\dot{V} = \Omega AV \; .$$

The solution of this equation can be represented by

(6.4) $$V(t) = \exp[(\Omega E-C)^{-1}\; \Omega A\; \frac{t}{\tau}]\; V(0) \; .$$

From this expression it is seen that Ω should be such that the real parts of the eigenvalues of the operator $(\Omega E-C)^{-1}\;\Omega A$ are largely negative. In that case we may expect a rapid convergence. In order to get information about the eigenvalues of $(\Omega E-C)^{-1}\;\Omega A$ we use the fact that the matrices E and A approximate differential operators as $h \to 0$. In our further considerations we restrict our analysis to the Dirichlet problem.

Locally we may write as $h \to 0$

$$
(6.5)\quad \begin{cases}
A \sim y_m \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} , \\
-E \sim \dfrac{y_m}{h^2} X_- + \dfrac{2}{l_m(l_m+l_{m+1})} Y_- \sim \\
\quad \sim \dfrac{y_m}{h^2} + \dfrac{2}{l_m(l_m+l_{m+1})} - \dfrac{y_m}{h} \dfrac{\partial}{\partial x} - \dfrac{2}{l_m+l_{m+1}} \dfrac{\partial}{\partial y} + \dots , \\
C = -2 \left(\dfrac{y_m}{h^2} + \dfrac{1}{l_m l_{m+1}}\right) .
\end{cases}
$$

Here $X_\pm$ and $Y_\pm$ denote the shift operators corresponding to x- and y-direction, respectively.

Next, we interprete the vector function $V(t)$ as being derived from a scalar function $V(x,y,t)$ by identifying the components of $V(t)$ with the values of $V(x_n,y_m,t)$. Substitution of (6.5) into (6.4) leads to a partial differential equation of the form

$$
(6.6)\quad aV_t + bV_{xt} + cV_{yt} = yV_{xx} + V_{yy} ,
$$

where on the line $y = y_m$

$$
(6.7)\quad \begin{cases}
a = \dfrac{\tau}{\Omega} \left[(2-\Omega) \dfrac{y_m}{h^2} - \Omega \dfrac{2}{l_m(l_m+l_{m+1})} + \dfrac{2}{l_m l_{m+1}}\right] , \\
b = \tau \dfrac{y_m}{h} , \\
c = \tau \dfrac{2}{l_m+l_{m+1}} .
\end{cases}
$$

Following an idea of Garabedian we introduce a new variable $z = z(x,y,t)$ such that equation (6.6) assumes the form

$$
(6.8)\quad FV_z + IV_{zz} = yV_{xx} + V_{yy} .
$$

A straightforward calculation reveals that the function $z(x,y,t)$ has to satisfy the differential equations

$$\text{(6.9)}\qquad \begin{aligned} bz_t - 2yz_x &= 0 , \\ cz_t - 2z_y &= 0 , \end{aligned}$$

and that the "friction" F and "inertia" I assume the form

$$\text{(6.10)}\qquad \begin{cases} F(x,y,t) = yz_{xx} + z_{yy} + az_t - c'(y)\, z_t , \\ I(x,y,t) = yz_x^2 + z_y^2 . \end{cases}$$

Again, we require that the solutions V(x,y,z) converge to zero as fast as possible.

We will try to satisfy this requirement by making the damping effect of F and I locally as large as possible. To that end we consider F and I as constants. Let e(x,y) be an eigenfunction of the Dirichlet problem and α its eigenvalues. Then,

$$\text{(6.11)}\qquad V = \exp\,[-\frac{F\pm\sqrt{F^2+4\alpha I}}{2I}\, z]\; e(x,y)$$

is a particular solution of (6.8). Assuming that α is real and negative, it is easily verified (see for instance reference [8], p.33) that (6.11) decreases the most rapidly to zero when

$$\text{(6.12)}\qquad F^2 + 4\alpha I = 0 .$$

Solution (6.11) is then given by

$$\text{(6.11')}\qquad V = \exp\,[-\sqrt{\frac{-\alpha}{I}}\, z]\; e(x,y) .$$

From this expression it follows that the largest damping of the general solution of (6.8) is obtained when (6.12) is satisfied for the absolute smallest eigenvalue, i.e.

$$\text{(6.12')}\qquad F^2 = 4|\alpha|_{min} I .$$

In order to determine F and I we have to solve equations (6.9) for z. Let us try a function z of the form

$$z = t + px + q(y) ,$$

where p is a constant and q(y) is an arbitrary function of y. Substitution in (6.9) yields

$$(6.13) \qquad p = \frac{b}{2y} \quad , \quad q(y) = \frac{1}{2} \int c(y)\, dy$$

and substitution in (6.20) yields

$$(6.10') \qquad \begin{aligned} F &= a - \frac{1}{2} c'(y) \\ I &= \frac{b^2}{4y} + \frac{c^2}{4} \quad . \end{aligned}$$

To simplify the calculations we apply the approximation

$$l_{m+1} \sim \frac{h}{\sqrt{y_m}} \qquad \text{as } h \to 0$$

in expressions (6.7) for a and c. It is easily verified that these expressions reduce to

$$(6.7') \qquad \left\{ \begin{aligned} a &\sim 2\tau \frac{2-\Omega}{\Omega} \frac{y}{h^2} \\ b &\sim \tau \frac{y}{h} \\ c &\sim \tau \frac{\sqrt{y}}{h} \end{aligned} \right. \quad , \text{ as } h \to 0$$

By these simplified formulae we obtain for F and I the expressions

$$(6.10'') \qquad \left\{ \begin{aligned} F &= \frac{\tau}{h} [2 \frac{2-\Omega}{\Omega} \frac{y}{h} - \frac{1}{4\sqrt{y}}] \quad , \\ I &= \frac{1}{2} \frac{\tau^2}{h^2} y \quad . \end{aligned} \right.$$

Relation (6.12') finally yields Ω as a function of y, i.e.

(6.14) $$\Omega = \frac{2}{1 + \frac{1}{2} h\left(\frac{1}{4y\sqrt{y}} + \sqrt{\frac{2|\alpha|_{min}}{y}}\right)} .$$

For numerical calculations it is convenient to reduce (6.14) by means of relation (4.7) to the form

(6.14') $$\Omega_m = \frac{2}{1 + \frac{1}{12m} + c\left[\frac{h^2}{m}\right]^{1/3}} ,$$

where c is a constant given by

$$c = \frac{\sqrt{|\alpha|_{min}}}{2^{1/6}\,3^{1/3}} \sim .617\ \sqrt{|\alpha|_{min}} .$$

In order to estimate the value of $|\alpha|_{min}$ we just derive an upper bound for $|\alpha|_{min}$. Consider the eigenvalue problem

$$\left(y \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)e = \alpha e,$$

where e satisfies homogeneous boundary conditions. Suppose that e(x,y) can be written as

$$e(x,y) = X(x)\ Y(y).$$

Then we have

$$y \frac{X''}{X} + \frac{Y''}{Y} = \alpha .$$

Hence,

$$\frac{X''}{X} = c ,$$

where c is a constant. Clearly, this equation is solved by

$$X = \sin n x\pi\ , \quad c = -m^2\pi^2, \quad n = \pm 1, \pm 2, \ldots .$$

The function Y satisfies the equation

$$T_n Y \equiv \left(\frac{d^2}{dy^2} - yn^2\pi^2\right) Y = \alpha Y ,$$

so that

$$|\alpha|_{min} \leq \operatorname*{Min}_n \frac{|(T_n\phi,\phi)|}{(\phi,\phi)} ,$$

where ϕ is an arbitrary function of y satisfying the boundary conditions and (,) denotes the inner product with respect to the interval $0 \leq y \leq 1$. For instance, the function $\phi = y(y-1)$ yields

$$|\alpha|_{min} \leq \operatorname*{Min}_n \frac{\int_0^1 (2-yn^2\pi^2)\, y(y-1)\, dy}{\int_0^1 y^2(y-1)^2\, dy} = \operatorname*{Min}_n \frac{5}{2}\,(n^2\pi^2-4) \sim 15.$$

We have done experiments with the SOR method in which Ω was given by (6.14) and $|\alpha|_{min} = 1,2,3,\ldots,15$. We found the largest rate of convergence for

$$|\alpha|_{min} = 12 .$$

In table 6.1 the results corresponding to this value of $|\alpha|_{min}$ are compared with the results obtained for a fixed value of Ω^+.

From this table it may be concluded that the methods are asymptotically equivalent. Only in the first iterations the variable Ω method leads to a slightly larger rate of convergence.

Table 6.1. Rates of convergence for fixed and varying values of Ω^+

h	k	Ω^+ fixed	Ω^+ variable	gainfactor
1/12	20	.43	.50	1.16
	40	.44	.53	1.20
	50	.46	.52	1.13
	60	.44	.44	1.00
	70	.38	.38	1.00
	80	.33	.33	1.00
	90	.29	.29	1.00
1/18	30	.31	.31	1.00
	60	.35	.36	1.03
	90	.29	.29	1.00

7. Numerical solution of the Tricomi problem

In the preceding sections we have analysed the Dirichlet problem for Tricomi's equation and found a formula for the optimal relaxation factor in the elliptic region. We now are in a position to investigate which relaxation factor should be used in the hyperbolic region.

We carried out a large number of experiments with varying values of the pair (Ω^+,Ω^-). In table 7.1 the results are given of the pair (Ω^+,Ω^-) which gave the largest rate of convergence.

Table 7.1. Rates of convergence for $\Omega^+ \neq \Omega^-$ in the elliptic, parabolic and hyperbolic region, respectively

h	k	Ω^+	Ω^-	rate of convergence			divergence
				ell.region	par.region	hyp.region	
30 60	1/6	1.35	1.15	.52 .45	.43 .42	.52 .44	$\Omega^- \geq 1.8$
40 60 90	1/12	1.54	1.10	.22 .20 .20	.12 .14 .15	.16 .17 .17	$\Omega^- \geq 1.4$
60 80 100 120	1/18	1.70	.85	.13 .12 .12 .12	.04 .05 .06 .07	.07 .08 .08 .08	$\Omega^- \geq 1.2$

As was already mentioned in section 5 the optimal relaxation factors differ largely (note that Ω^+ is close to the optimal value for the Dirichlet problem). Hence, it is expected that iterating with $\Omega^+ = \Omega^-$ will give a lower rate of convergence. In table 7.2 results are given for the optimal value of a fixed relaxation factor.

Table 7.2. Rates of convergence for fixed Ω in the elliptic, parabolic and hyperbolic region, respectively

h	k	Ω	rate of convergence			divergence
			ell.region	par.region	hyp.region	
1/6	30 60	1.28	.48 .45	.40 .43	.49 .46	$\Omega \geq 1.7$
1/12	40 60 90	1.26	.17 .15 .14	.08 .10 .11	.12 .12 .12	$\Omega \geq 1.4$
1/18	60 80 100 120	1.20	.10 .09 .08 .07	.04 .04 .04 .04	.06 .06 .06 .06	$\Omega \geq 1.3$

We see that the greater the difference between the optimal values of Ω^+ and Ω^- given in table 7.1, the lower the rate of convergence when Ω is kept fixed.

Finally, the results are given when the original Fillippov scheme (method of Gauss-Seidel) is applied to the Tricomi problem.

Table 7.3. Rates of convergence for Gauss-Seidel's method in the elliptic, parabolic and hyperbolic region, respectively.

h	k	rate of convergence		
		ell.region	par.region	hyp.region
1/6	30	.32	.27	.30
	60	.30	.28	.29
1/12	40	.14	.07	.10
	60	.11	.07	.09
	90	.10	.07	.08
1/18	60	.09	.03	.06
	80	.08	.03	.05
	100	.07	.03	.05
	120	.06	.03	.04

8. References

[1] Bergman, S.: Application of integral operators to the theory of partial differential equations with singular coefficients. Arch. Rat. Mech. Anal. (1962).

[2] Chia-Kun Chu: Type insensitive finite difference schemes. Thesis New York (1958).

[3] Fillippov, A.F.: On the application of the method of finite differences to the solution of the problem of Tricomi. Izv. Akad. Nauk SSSR. Ser. Math. 21, 73-88 (1957) (Russian).

[4] Frankel, S.P.: Convergence rates of iterative treatments of partial differential equations. Math. Tables Aids Comput. 4, 65-75 (1950).

[5] Guderly and Yoshihara: The flow over a wedge profile at Mach Number 1. J. Aeronautical Sciences, 17 (1950).

[6] Houwen, P.J. van der: Some analytical aspects of the Tricomi problem. Report TW 102, Math. Centrum, Amsterdam (1966).

[7] Houwen, P.J. van der: On the acceleration of Richardson's method II. Numerical aspects. Report TW 107, Math. Centrum, Amsterdam (1967).

[8] Coolen, T.M.T. and P.J. van der Houwen: On the acceleration of Richardson's method IV. A non-symmetrical case. Report TW 109, Math. Centrum, Amsterdam (1968).

[9] Levy, D.: A numerical scheme for solving a boundary value problem for the Tricomi equation. AEC Computing and Appl. Math. Center NYO, 7979 (1958).

[10] Morawetz, C.A.: A weak solution for a system of equations of elliptic-hyperbolic type. Comm. Pure Appl. Math. (1958).

[11] Ovsyannikov, L.V.: On Tricomi's problem in a class of generalized solutions of the Euler-Darboux equation. Doklady Akad. Nauk SSSR (N.S.) 91, 457-460 (1953) (Russian).

[12] Varga, R.: Matrix iterative analysis.
Prentice-Hall, Inc. Englewood Cliffs, New Yersey (1962).

[13] Vicenti, W.G. and C.B. Wagoner: Transonic flow past a wedge profile with detached bow wave.
NACA Rep. no. 1095 (1952).

[14] Young, D.M.: Iterative methods for solving partial difference equations of elliptic type.
Trans. Amer. Math. Soc. 76, 92-111 (1954).