Journal of Computational Mathematics, Vol.10, No.1, 1992, 57–76.

ANALYSIS OF FLOW DIRECTED ITERATIONS*

Han Hou-de
(*Department of Applied Mathematics, Tsinghua University, Beijing, China*)
V.P. Il'in
(*Computing Center, Siberian Division of the Academy of Sciences, Novosibirsk, USSR.*)
R.B.Kellogg
(*Institute for Phys. Sci. and Tech., University of Maryland, U.S.A.*)
Yuan Wei
(*Department of Applied Mathematics, Tsinghua University, Beijing, China*)

Abstract

Iterative Methods are studied for the solution of difference schemes for convection dominated flow problems.

§1. Introduction

The numerical solution of convection diffusion flow problems is of considerable difficulty. It is well known that when the diffusion coefficient is small (the case of convection dominated flow), it is hard to obtain accurate difference schemes; the presence of rapid transitions, or boundary layers, in the solution severely degrades the accuracy of the approximate solution. One may ask whether difficulties are also encountered in the numerical solution of the difference equations when the diffusion coefficient is small. In this paper we consider some difference approximations to the convection diffusion equation and we treat block Gauss Seidel iterations for the solution of these problems. We study the effect of the partitioning and ordering of the unknowns on the convergence of the Gauss-Seidel iterations. We find that, for convection dominated flow problems, the spectral radius of the iteration matrix is not an appropriate indicator of the convergence properties of the method; it is better to use a norm of the iteration matrix. Also, we find that sweeping the mesh in the direction of the underlying flow enhances the convergence of the Gauss Seidel iterations. In one dimension it is not hard to devise an algorithm to implement this idea. In two dimensions, we give a general procedure to automate the partitioning and ordering phase of the solution process. The general procedure is described using the graph of the matrix.

§2 contains some remarks about the Gauss Seidel method. In §3 we discuss the one dimensional problem. For the basic upwind difference scheme on a uniform

* Received October 18, 1988..

mesh we find that if the unknowns are swept in the direction of flow, the norm of the iteration matrix satisfies the inequality $\|N\| \leq c\varepsilon/h^s$ where $s = 2$ or 3. Thus while the iterations in any event converge, if $\varepsilon << h$, the iterations converge very quickly. Since, with $\varepsilon << h$, the difference equations are basically solving the reduced problem, which is a first order equation, it may not be too surprising that the convergence is fast. However we find that this fast convergence also holds for difference schemes that are especially adapted to the solution of the convection diffusion problem, such as the exponential scheme of Southwell, Allen, and A.M. Il'in and a difference scheme of Ervin and Layton that has been found to provide good resolution of interior layers. We have performed numerical tests for these schemes, and also for discretizations with a refined mesh that is designed to capture the boundary layer. In the latter case, flow directed iterations do not perform as well, but they are better than iterating with other orderings of the unknowns. §4 deals with the basic upwind scheme in two dimensions. Here we find the same inequality for $\|N\|$ if flow directed iterations are used. The difficulty lies in ordering the nuknowns to implement flow directed iterations. We find an algorithm that solves this problem. Tests of this algorithm, and of other ordering procedures for two dimensional problems, are not given in this paper.

The conclusion of this study is that flow directed iterations may be a good way to solve the discrete equations arising from modelling convection dominated flow. If the boundary or interior layers are captured by refining the mesh, the convergence properties of the method are not so favorable. Further work on flow directed iterations must be done to deal with refined meshes, and to develop the method for nonlinear problems and for the Navier Stokes equations, which contain a continuity equation as well as equations of convection diffusion type.

In [8], Strikwerda has considered SOR methods for the iterative solution of convection diffusion problems. We describe his approach in §4. Some recent work of Goldstein on the use of preconditioned conjugate gradients for solving convection diffusion equations is given in [6].

§2. Some properties of the Gauss Seidel Method

We recall the Gauss Seidel method for solving a linear system $Au = f$. Let us write $A = D - L - U$, where, typically, D is a nonsingular diagonal or block diagonal matrix and, for some permutation matrix, P, PLP^{-1} and PUP^{-1} are respectively lower and upper block triangular. The Gauss Seidel iterations may be written $Du^{k+1} = Lu^{k+1} + Uu^k + f$ or, solving for $u^{k+1}, u^{k+1} = Nu^k + (D-L)^{-1}f$, where $N = (I - D^{-1}L)^{-1}D^{-1}U$ is the Gauss Seidel iteration matrix. Since $D^{-1}L$ is typically block triangular, $D^{-1}L$ is usually nilpotent. We say that $D^{-1}L$ is nilpotent of index m if $(D^{-1}L)^m = 0$, and if m is the smallest index for which this holds. We shall frequently use the following simple lemma to estimate $\|N\|$.

Lemma 1. *If D is invertible, if $D^{-1}L$ is nilpotent of index k, and if $\|D^{-1}L\| \leq \alpha < 1$ and $\|D^{-1}U\| \leq \beta$ then*

$$\|N\| \leq \frac{1-\alpha^k}{1-\alpha}\beta \leq \frac{\beta}{1-\alpha}.$$

Proof. Using the geometric series and the nilpotency of $D^{-1}L$,

$$\|N\| \leq \|(I - D^{-1}L)^{-1}\| \cdot \|D^{-1}U\| \leq (1 + \alpha + \cdots + \alpha^{k-1})\beta = \frac{1-\alpha^k}{1-\alpha}\beta.$$

This result is valid for an arbitrary matrix norm, but here and in the following we use only the ∞ norm for vectors, and the corresponding matrix norm, defined for a matrix $A = [a_{ij}]$ by

$$\|A\| = \|A\|_\infty = \max_i \left\{ \sum_j |a_{ij}| \right\}.$$

If $\alpha + \beta = 1$ and $\beta \to 0$, then $\|N\| \to 0$ and we may obtain a very accurate solution after one iteration. Note that for "reverse" Gauss Seidel iterations, defined by $Du^{k+1} = Lu^k + Uu^{k+1} + f$, the iteration matrix is $\tilde{N} = (I - D^{-1}U)^{-1}D^{-1}L$, and $\|\tilde{N}\| \to 1$ as $\beta \to 0$.

It is interesting to note that this analysis of the convergence for small β does not follow from spectral radius considerations. If the matrix A is block tridiagonal (or of "type A" [9]), both N and $\tilde{N}$ have the same spectral radius, namely, $\rho(N) = \rho(\tilde{N}) = \rho(N_J)^{1/2}$, where $N_J = D^{-1}(L + U)$ is the Jacob iteration matrix. By a similarity transformation it is easy to derive the estimate $\rho(N) \leq 2\sqrt{\alpha\beta}/(\alpha + \beta)$. Hence $\rho(N_J) \to 0$ as $\beta \to 0$. In this case, the Gauss Seidel method, the Jacobi method, and the "reverse" Gauss Seidel method, have a fast asymptotic rate of convergence, in the sense that $R_\infty = \lim(\|y^k\|/\|y^0\|)^{1/k} = \rho$ is small, where $y^k = u - u^k$ is the error after k iterations. However, the error reduction for the last two methods, as measured in a norm of interest, is not small for a moderate number of iterations. We conclude that the spectral radius of the iteration matrix is not a good indictor of the convergence properties of the Gauss Seidel method when β is small. We shall therefore estimate norms, not spectral radius, in our analysis (Conditions which guarantee that the spectral radius of a matrix N equals either $\|N\|_\infty$ of $\|N\|_1$ have recently been given by W.T. Tong [10].)

Another conclusion of this analysis is that it is not fruitful to use overrelaxed iterations for these problems. To see this, again suppose that A is tridiagonal or has property A. Then as $\beta \to 0, \rho(N_J) \to 0$ and the optimal SOR parameter $\omega_b \to 1$. Thus, for small β, optimal overrelaxation is basically the Gauss Seidel method. In §4 we will seek orderings of the unknowns which produce a small value of β. These orderings will result in permuted matrices A which do not necessarily satisfy property A, but which nevertheless give rise to very rapidly convergent iterations. As will be seen, the parameter β basically corresponds to the value of the diffusion in a singularly perturbed convection diffusion equation.

§3. Some Examples in One Dimension

We start with the one dimensional problem

$$-\varepsilon u'' + p(x)u' + r(x)u = f(x), \quad 0 < x < 1, \tag{3.1a}$$

$$u(0) = u(1) = 0. \tag{3.1b}$$

The equation may be interpreted as representing the diffusion and convection of a solute in a one dimensional medium. $u(x)$ denotes the concentration of the solute, ε denotes the diffusion coefficient and $p(x)$ denotes the rate of flow of the fluid. If $p(x) > 0$ (< 0) the flow is to the right (left). The term ru represents an absorption of solute; we assume that $r(x) \geq 0$. We approximate (3.1) on a uniform mesh of size $h = 1/(n+1)$. We write the upwind difference approximation as

$$-b_i u_{i-1} + a_i u_i - c_i u_{i+1} = f_i, \quad 1 \leq i \leq n, \quad u_0 = u_{n+1} = 0. \tag{3.2}$$

The coefficients a_i, b_i, c_i are given by

$$b_i = \frac{\varepsilon}{h^2} + \frac{1}{2h}(|p_i| + p_i), \quad c_i = \frac{\varepsilon}{h^2} + \frac{1}{2h}(|p_i| - p_i), \quad a_i = b_i + c_i + r_i h, \tag{3.3}$$

where $p_i = p(x_i), r_i = r(x_i)$. The matrix $A \in R^{n\times n}$ is the tridiagonal matrix corresponding to the difference equation (3.2). We let A_0 be the corresponding matrix with the r_i set equal to 0.

We first consider the case when $p(x) > 0$ on $[0,1]$. We decompose A into its lower triangular, diagonal, and upper triangular parts, $A = D - L - U$. In this case, $D^{-1}L$ and $D^{-1}U$ are respectively strictly lower and upper triangular matrices with one nonzero diagonal adjacent to the main diagonal, whose entries α_i and β_i are given respectively by the formulas

$$\alpha_i = \frac{\varepsilon + |p_i|h}{2\varepsilon + |p_i|h + r_i h^2}, \quad \beta_i = \frac{\varepsilon}{2\varepsilon + |p_i|h + r_i h^2}, \tag{3.4}$$

(In this case it is not necessary to write $|p_i|$; we do so in order that (3.4) may be used in later situations.) Writing $A_0 = D_0 - L_0 - U_0$, we see from (3.4) that $D^{-1}L \leq D_0^{-1}L_0$, and $D^{-1}U \leq D_0^{-1}U_0$, where the inequalities are meant in an entrywise sence. It then follows from the geometric series that $N \leq N_0$, where $N_0 = (I - D_0^{-1}L_0)^{-1}D_0^{-1}U_0$. We therefore have $\|N\| \leq \|N_0\|$. Let $\hat{p} = \max_i |p_i|, \check{p} = \min_i |p_i|$, with a similar notation for $\check{r}$ and, in §4, $\check{q}, \hat{q}$. From (3.4) we see that

$$\|D_0^{-1}U_0\| = \frac{\varepsilon}{2\varepsilon + \check{p}h}, \tag{3.5a}$$

$$\|D_0^{-1}L_0\| = \frac{\varepsilon + \hat{p}h}{2\varepsilon + \hat{p}h}, \tag{3.5b}$$

so

$$1 - \|D_0^{-1}L_0\| = \frac{\varepsilon}{2\varepsilon + \hat{p}h} \le \frac{\varepsilon}{\hat{p}h}.$$

We use these formulas and the lemma to estimate $\|N_0\|$. Since $n(1-\tau)^{n-1} < n$ on $(0,1]$, we may integrate both sides to get

$$1 - (1-\tau)^n < n\tau, \quad \text{for } 0 < \tau \le 1. \tag{3.6}$$

We use (3.6) with $\tau = \varepsilon/\hat{p}h$ to obtain

$$1 - \|D_0^{-1}L_0\|^n \le 1 - (1-\tau)^n \le n\varepsilon/\hat{p}h \text{ if } \varepsilon \le \hat{p}h.$$

Hence, using the lemma with $k = n$,

$$\|N_0\| \le \frac{n\varepsilon}{\hat{p}h} \cdot \frac{2\varepsilon + \hat{p}h}{h} \cdot \frac{\varepsilon}{2\varepsilon + \check{p}h} < \frac{n\varepsilon}{\check{p}h}.$$

We thus obtain

$$\|N\| \le \|N_0\| \le \frac{\varepsilon}{\check{p}h^2}. \tag{3.7}$$

If ε is sufficiently small, $\|N\| < 1$. We conclude that while the Gauss Seidel iterations are, in any event, convergent, in the parameter range where $\varepsilon < 2\check{p}h^2$, the error reduction per iteration as measured in the ∞ norm is $< 1/2$.

A similar conclusion is obtained if an "exponential" difference scheme is used. The Southwell-Allen-A. M. Il'in scheme is given by (3.2), where the coefficients satisfy

$$b_i = \frac{p_i}{2h}\coth\frac{p_i h}{2\varepsilon} + \frac{p_i}{2h}, \quad c_i = \frac{p_i}{2h}\coth\frac{p_i h}{2\varepsilon} - \frac{p_i}{2h}, \quad a_i = b_i + c_i + r_i, \tag{3.3'}$$

instead of (3.3). (See, for example, [7]). The matrix A is of "positive type', that is, $b_i > 0, c_i > 0$. Suppose $p(x) > 0$ on $[0,1]$. Decompose A into its diagonal. lower triangular, and upper triangular parts, $A = D - L - U$, and in a similar way write $A_0 = D_0 - L_0 - U_0$. As above, we again find that $D^{-1}L, D^{-1}U$, and N are majorized by $D_0^{-1}L_0, D_0^{-1}U_0$ and N_0 respectively. The matrices $D_0^{-1}L_0$ and $D^{-1}U_0$ are respectively lower and upper bidiagonal matrices with zero diagonal and off diagonal entries α_i^0 and β_i^0 given by the formulas

$$\alpha_i^0 = \frac{1}{1 + \exp(-p_i h/\varepsilon)}, \quad \beta_i^0 = \frac{\exp(-p_i h/\varepsilon)}{1 + \exp(-p_i h/\varepsilon)}. \tag{3.4'}$$

calculating the norms, we obtain

$$\|D_0^{-1}L_0\| = \frac{1}{1 + \exp(-\hat{p}_i h/\varepsilon)}, \quad \|D_0^{-1}U_0\| = \frac{\exp(-\check{p}_i h/\varepsilon)}{1 + \exp(-\check{p}_i h/\varepsilon)},$$

so

$$1 - \|D_0^{-1}L_0\| = \frac{\exp(-\hat{p}h/\varepsilon)}{1+\exp(-\hat{p}h/\varepsilon)} \leq \exp(-\hat{p}h/\varepsilon).$$

We use (3.6) with $\tau = \exp(-\hat{p}h/\varepsilon)$ to obtain

$$1 - \|D_0^{-1}L_0\|^n \leq n\exp(-\hat{p}h/\varepsilon).$$

Hence, using the lemma,

$$\|N_0\| \leq n\exp(-\hat{p}h/\varepsilon)\cdot\frac{1+\exp(-\hat{p}h/\varepsilon}{\exp(-\hat{p}h/\varepsilon)}\cdot\frac{\exp(-\check{p}h/\varepsilon}{1+\exp(-\check{p}h/\varepsilon)} \leq 2n\exp(-\check{p}h/\varepsilon).$$

Since $\exp(-\tau) \leq \tau^{-1}$ for $\tau > 0$, we obtain $\|N_0\| \leq 2n\varepsilon/\check{p}h$. Thus, $\|N\| \leq 2\varepsilon/\check{p}h^2$, and we again conclude that in an appropriate parameter range, the error reduction per iteration is $< 1/2$.

We have also considered a difference scheme of Ervin and Layton [4, 5] which does not require the evaluation of exponentials, which satisfies an $O(h^2)$ error estimate for problems without turning points when $\varepsilon << h$, and which in practice reproduces sharp interior layers in many problems. We have found that flow directed iterations work very well for the schemes of Ervin and Layton.

We next consider several cases when p has a change of sign. It is convenient to suppose here that $\check{r} > 0$. Suppose first that p is a decreasing function with $p(0) > 0, p(1) < 0$ and with $p'(x) < 0$ in $[0,\ 1]$. In this case, the zero x^* of p corresponds to a sink; the flow is directed towards x^* on both sides of x^*. Motivated by the preceding discussion, we sweep the mesh in the direction of the flow. To describe the iterative process, suppose that x^* lies in the interval (x_k, x_{k+1}). Then we start at $i = 1$ and perform a Gauss Seidel sweep, successively solving for u_i for $i = 1, 2, \cdots, k$. Next we solve the n-th equation for u_n, and successively solve the $(n-1)$-st equation for $u_{n-1}, \cdots$, down to solving to the $(k+1)$-st equation for u_{k+1}. An analysis of the iteration matrix may be given following the proof of Theorem 2 in §4. We obtain

$$\|N\| \leq \frac{1(\hat{p}+\check{r})}{\check{r}}\cdot\frac{\varepsilon}{h^3}. \tag{3.11}$$

We conclude that while the iterations are in any event convergent, in a suitable parameter range the error reduction per iteration as measured in the ∞ norm is $< 1/2$.

Finally, suppose p is an increasing function with $p(0) < 0, p(1) > 0$. and with $p'(x) > 0$ in $[0,\ 1]$. In this case, the analysis is more difficult. The zero x^* of p corresponds to a source; the flow is directed away from x^* on both sides of x^*. Near each end of the interval the flow is out of the interval, so the sweep must start inside the interval. Suppose that $x^* \in (x_k, x_{k+1})$. The iterative method will be a block Gauss Seidel method, where the first block has size 2 and the remaining blocks have size 1. In a single iterative sweep, we first solve simultaneously the k-th, and $(k+1)$-st equations for v_k and v_{k+1}, and we then update the remaining unknowns in the order $v_{k-1}, \cdots, v_1, v_{k+2}, \cdots, v_n$. With this iterative strategy, the matrix D consists of the principal 2×2 submatrix of A defined by rows and columns k and $k+1$, and the remaining diagonal entries of A. The matrices L and U are defined similarly. An analysis of the iteration matrix along the lines of Theorem 2 again yields (3.11). Again we obtain rapid convergence of the Gauss Seidel iterations when ε and h lie in a certain range.

It is possible to avoid the use of a 2×2 block in the case of a source type turning point by using an alternate ordering of the mesh points. We call this ordering the FDPI ordering of the mesh points. Further examples of special orderings will be given in §5. If $p(x)$ is as in the preceding paragraph, with a single zero at a point $x^* \in (x_k, x_{k+1})$, in the FDPI ordering we sweep the mesh in the direction left to right, going from $i = k+1$ to $i = n$, and then we sweep the mesh in the direction right to let, going from $i = k$ to $i = 1$. With this ordering, the flow directed pattern is broken at a single point, the point $i = k+1$. The coefficient linking u_{k+1} with u_k in the $(k+1)$-st equation is an entry of the matrix U, but is not of order ε. In the case of a general function $p(x)$, the ordering is described in terms of the set $\mathcal{N} = \{1, \cdots, n-1\}$. Let $\mathcal{N} = \mathcal{N}_E \cup \mathcal{N}_W$, where

$$\mathcal{N}_E = \{i \in \mathcal{N} \ : \ p_i \geq 0\}; \quad \mathcal{N}_W = \{i \in \mathcal{N} \ : \ p_i < 0\}.$$

The FDPI ordering consists of the points of $\mathcal{N}_E$ arranged from left to right, followed by the points of $\mathcal{N}_W$ arranged from right to left. The following theorem gives an estimate for the iteration matrix of the method. The theorem suggests that, while we cannot expect extremely rapid convergence for $\varepsilon << h$, we can in any event expect convergence behavior that does not degrade for small h, provided that $\varepsilon << h$.

Theorem 1. *Let N be the iteration matrix of FDPI. Suppose* $\max |p'(x)| \leq \hat{p}'$, *and let the bounds $\hat{p}', \hat{p}$, and $\check{r}$ be chosen so that $\check{r} < \hat{p}', \check{r} < \hat{p}$. Then*

$$\|N\| \leq \frac{\hat{p}'}{\hat{p}' + \check{r}} + \frac{\hat{p}}{\check{r}^2} \cdot \frac{\varepsilon}{h^3}.$$

Proof. Let $\mathcal{N}_0 = \{i : 1 \leq i \leq n-1, p_i \geq 0, p_{i-1} < 0\}$. Then for $i \in \mathcal{N}_0$, both $i+1$ and $i-1$ come after i in the FDPI ordering. so both b_i and c_i are entries on the row of U corresponding to the mesh point i. Let $U = U_0 + U_1$, where U_0 contains the entries b_i of U, for $i \in \mathcal{N}_0$, and where U_1 contains the remaining entries of U. Let $N_l = (D-L)^{-1}U_l, l = 0, 1$, so $N = N_0 + N_1$ is the FDPI iteration matrix. Each row of $D^{-1}L$ contains at most one non-zero entry, so

$$\|D^{-1}L\| \leq \max_i \frac{\varepsilon + |p_i|h}{2\varepsilon + |p_i|h + r_i h^2} \leq \frac{\varepsilon + \hat{p}h}{2\varepsilon + \hat{p}h + \check{r}h^2}.$$

This expression is a decreasing function of ε provided $\hat{p} > \check{r}$. Hence $\|D^{-1}L\| \leq \hat{p}/(\hat{p} + \check{r}h)$, so $1 - \|D^{-1}L\| \geq \check{r}h/\hat{p}$, so $\|(I - D^{-1}L)^{-1}\| \leq \hat{p}/\check{r}h$. Each row of U_1 contains at most one non-zero entry, so

$$\|D^{-1}U_1\| \leq \max_i \frac{\varepsilon}{2\varepsilon + |p_i|h + r_i h^2} \leq \frac{\varepsilon}{\check{r}h^2}.$$

Hence $\|N_1\| \leq (\varepsilon\hat{p}/(\check{r}^2h^3))$. To estimate N_0 we use the formula $N_0 = D^{-1}U_0 + D^{-1}LU_0 + \cdots$. $D^{-1}U_0$, and hence $(D^{-1}L)^k D^{-1}U_0$, has entries only on those columns corresponding to nodes $j \in \mathcal{N}_0$. Each entry of $(D^{-1}L)^k U_0$ corresponds to a path of length k in $\mathcal{G}(L)$ that terminates in a point $j \in \mathcal{N}_0$, and the value of the entry is the product of the entries of $D^{-1}L$ corresponding to links on the path, multiplied by b_j. Since there is at most one path in $\mathcal{G}(L)$ leading from any point i to a point $j \in \mathcal{N}_0$, each row of N_0 contains at most one non-zero entry. Hence

$$\|N_0\| \leq \max\left\{\frac{\varepsilon + |p_j|h}{2\varepsilon + |p_j|h + r_j h^2} : j \in \mathcal{N}_0\right\}.$$

If $j \in \mathcal{N}_0, p(x)$ vanishes within a distance h of jh, so $|p_j| \leq \hat{p}'h$. Hence

$$\|N_0\| \leq \frac{\varepsilon + \hat{p}'h^2}{2\varepsilon + \hat{p}'h^2 + \check{r}h^2}.$$

This quantity is a decreasing function of ε if $\check{r} < \hat{p}'$. Hence, assuming this, $\|N_0\| \leq \hat{p}'/(\hat{p}' + \check{r})$.

If p has a number of zeros, we may construct an iterative strategy based on a combination of the above methods. We obtain a norm estimate for the iteration matrix similar to (3.11). If the exponential scheme (3.2), or the Ervin-Layton scheme, is used on problems with turning points, similar conclusions have been found with regard to the behavior of flow directed iterations.

§4. Two Dimensional Problems

We consider the convection diffusion equation in the unit square $\Omega = (0,1) \times (0,1)$:

$$-\varepsilon(u_{xx} + u_{yy}) + pu_x + qu_y + ru = f, \quad (x,y) \in \Omega,, \tag{4.1a}$$

$$u(x,y) = 0, \quad (x,y) \in \partial\Omega. \tag{4.1b}$$

The equation may be interpreted as representing the diffusion and convection of a solute in a two dimensional medium. The vector (p,q) represents the flow of the solvent.

In [8]. Strikwerda considers the use of SOR iterations to solve the convection diffusion equation on a square Ω. The approach is somewhat different than ours, and involves a refined mesh that is chosen specifically to capture the boundary layers of the problem. By using a refined mapping, a refined mesh is introduced along the sides of Ω where the boundary layers of the solution occur. A discretization of the problem on the variable mesh is used. The overrelaxation parameter is taken to be mesh dependent, and a heuristic formula is developed for the choice of this parameter. The formula is based on an analysis of the difference equation with coefficients frozen at the mesh point. The SOR sweeps are made in the usual order; that is, with out regard to the direction of flow. Heuristic estimates for the spectral radius of the iteration matrix are found that are independent of ε. Strikwerda's approach is complementary to ours. It would be interesting to combine his heuristic formulas for the overrelaxation parameter with our use of flow directed sweeps of the mesh.

We approximate (4.1) on a uniform mesh of size $h = 1/(n+1)$. We shall consider difference approximations to (4.1) of the form

$$-a_{i,j}u_{i-1,j} - b_{i,j}u_{i,j-1} - c_{i,j}u_{i+1,j} - d_{i,j,}u_{i,j+1} + e_{i,j}u_{i,j} = h^2 f_{i,j}, \quad 1 \le i,j \le n, \quad (4.2)$$

where we set $u_{i,j} = 0$ if i or j is o or n. If we use an upwind difference approximation, for which each of the first derivatives u_x and u_y is differenced according to the sign of the coefficient p and q respectively, the coefficients in (4.2) are given by the formulas

$$a_{i,j} = \varepsilon + \frac{h}{2}(|p_{i,j}| + p_{i,j}), \quad b_{i,j} = \varepsilon + \frac{h}{2}(|q_{i,j}| + q_{i,j}), \quad c_{i,j} = \varepsilon + \frac{h}{2}(|p_{i,j}| - p_{i,j}),$$
$$d_{i,j} = \varepsilon + \frac{h}{2}(|q_{i,j}| - q_{i,j}), \quad e_{i,j} = 4\varepsilon + h(|p_{i,j}| + |q_{i,j}) + r_{i,j}h^2.$$

From the examples in §3 we conclude that it is wise to use Gauss Seidel iterations for which the direction of the iteration conforms to the direction of the flow. We first give an example to illustrate the type of iterations that result from this principle, and we then formulate a general strategy for constructing a Gauss Seidel iteration process from a prescribed flow vector (p, q). In our analysis we shall assume, for simplicity, that $r(x, y) > 0$. Some of the results can be extended to the case $r = 0$.

In our example, we assume that $p(x, y) \ge \check{p} > 0$ in Ω. In this case, the flow is directed from left to right. but the flow may go either up or down. It is appropriate to consider a block Gauss Seidel method, with the blocks corresponding to the lines $i = 1, 2, \cdots, n$. Decomposing the coefficient matrix $A = D - L - U$ in accordance with this iteration strategy, we see that D, L and U are defined by the equations

$$(Dz)_{i,j} = e_{i,j}z_{i,j} - b_{i,j}z_{i,j-1} - d_{i,j}z_{i,j+1},$$
$$(Lz)_{i,j} = a_{i,j}z_{i-1,j}, \quad (Uz)_{i,j} = c_{i,j}z_{i+1,j}.$$

Let $E = \operatorname{diag}(e_{i,j})$, and write $D = E - B$. Then $D - L = E[I - E^{-1}(L+B)]$. Since

$$\|E^{-1}(L+B)\| = \max_{i,j} \frac{2\varepsilon + p_{i,j}h + |q_{i,j}|h}{4\varepsilon + p_{i,j}h + |q_{i,j}|h + r_{i,j}h^2} \le \frac{2\varepsilon + \hat{p}h + \hat{q}h}{4\varepsilon + \hat{p}h + \hat{q}h + \hat{r}h^2} < 1,$$

we have

$$\|(D-L)^{-1}\| \le \frac{\|E^{-1}\|}{1-\|E^{-1}(L+B)\|} \le \frac{4\varepsilon+\hat{p}h+\hat{q}h+\hat{r}h^2}{[2\varepsilon+\check{r}h^2][4\varepsilon+\check{p}h+\check{q}h+\check{r}h^2]}.$$

The quantity $(4\varepsilon+\hat{p}h+\hat{q}h+\hat{r}h^2)/(4\varepsilon+\check{p}h+\check{q}h+\check{r}h^2)$ is a decreasing function of $4\varepsilon+\check{r}h^2$. Hence, replacing this quantity by 0, we get

$$\|(D-L)^{-1}\| \le \frac{1}{2\varepsilon+\check{r}h^2}\cdot\frac{\hat{p}+\hat{q}}{\check{p}+\check{q}} \le \frac{\hat{p}+\hat{q}}{\check{p}\check{r}}\cdot h^{-2}.$$

Since $N=(D-L)^{-1}U$, and since, in this case, $c_{i,j}=\varepsilon$, so $\|U\|=\varepsilon$, we obtain

$$\|N\| \le \frac{\hat{p}+\hat{q}}{\check{p}\check{r}}\cdot\frac{\varepsilon}{h^2}. \tag{4.3}$$

Hence, for mesh spacings h which satisfy $\varepsilon < bh^2/2$, where $b < \check{p}\check{r}/(\hat{p}+\hat{q})$, the error per iteration as measured in the ∞ norm is $< 1/2$.

We now consider the case of a general flow vector (p,q). It is convenient to introduce some notation. We let $\mathcal{N}$ denote the set of mesh points $P=(ih,jh), 1 \le i,j \le n$. We say that two mesh points $P=(ih,jh)$ and $Q=(kh,lh)$ are neighbors if $|i-k|+|j-l|=1$. If P and Q are neighbors, we say that the ordered pair (p,Q) is a β link if the coefficient of u_Q in the difference equation (4.2) for u_P is $-\varepsilon$. If (P,Q) is not a β link, we say that (P,Q) is an α link. If (P,Q) is an α link, the flow vector (p_{ij},q_{ij}) at P points away from Q in the following sense: if P lies to the left of $Q(i=k-1)$, then $p_{ij}<0$; if P lies to the right of Q, then $p_{ij}>0$; if P lies below Q, then $q_{ij}<0$; and if P lies above Q, then $q_{ij}>0$. Of course (P,Q) and (Q,P) may not be links of the same type. We define a directed graph, or digraph, $\mathcal{G}_0$, as follows. The node set of $\mathcal{G}_0$ is the set $\mathcal{N}$ of mesh point. If $P,Q\in\mathcal{N}$, then (P,Q) is a directed link of $\mathcal{G}_0$ provided (P,Q) is an α link.

By a partition of the node set $\mathcal{N}$, we mean a collection of disjoint subsets of $\mathcal{N}$ whose union is $\mathcal{N}$. By an ordered partitions of $\mathcal{N}$ we mean a partition of $\mathcal{N}$ with the subsets written in a definite order. Each ordered partition $\mathcal{N}_1,\cdots,\mathcal{N}_m$ of $\mathcal{N}$ gives rise to a block Gauss Seidel method in the following manner. We reorder and partition a vector v of unknowns as $v=[v^1,\cdots,v^m]$ to conform with the given ordered partition of $\mathcal{N}$, and in a similar manner, we permute and partition the rows and columns of the coefficient matrix A. Writing $A=D-L-U$, with $D_i, L_{i,j}$, and $U_{i,j}$ the corresponding permuted and partitioned submatrices, $1\le i,j\le m$, we may write the i-th set of difference equations as

$$D_iv^i-\sum_{j<i}L_{i,j}v^j-\sum_{j>i}U_{i,j}v^j=h^2f^i. \tag{4.4}$$

The block Gauss Seidel method then consists in first solving (4.4) with $i=1$ for v^1, then updating v^1 and solving (4.4) with $i=2$ for $v^2,\cdots,$. The corresponding iteration matrix is $N=(D-L)^{-1}U$.

We say that an ordered partition $\mathcal{N}_1, \cdots, \mathcal{N}_m$ of $\mathcal{N}$ is admissible provided the following condition is satisfied: if $k < l$ if $P \in \mathcal{N}_k$ and $Q \in \mathcal{N}_l$ and if P and Q are neighbors, then (P, Q) is a β link. The following theorem shows that admissible ordered partitions give rise to rapidly convergent block Gauss Seidel methods.

Theorem 2. *If $\mathcal{N}_1, \cdots, \mathcal{N}_m$ is an admissible partition of $\mathcal{N}$, and if N is the corresponding Gauss Seidel iteration matrix, then*

$$\|N\| \leq 3\frac{\check{r} + \hat{p} + \hat{q}}{\check{r}^2} \cdot \frac{\varepsilon}{h^3}. \tag{4.5}$$

Proof. Write $D = E - B$, where E is the diagonal matrix of D, and hence the diagonal matrix of A. Then $D - L = E - L - B = E[I - E^{-1}(L + B)]$. Since $L + B \leq L + U + B$ in the entrywise sense, we have

$$\|E^{-1}(L + B)\| \leq \|E^{-1}(L + B + U)\| \leq \max_{i,j} \frac{4\varepsilon + |p_{i,j}|h + |q_{i,j}|h}{4\varepsilon + |p_{i,j}|h + |q_{i,j}|h + r_{i,j}h^2}.$$

The ratio is an increasing function of the quantity $|p_{i,j}|h + |q_{i,j}|h$, which leads to the estimate

$$\|E^{-1}(L + B)\| \leq \frac{4\varepsilon + \hat{p}h + \hat{q}h}{4\varepsilon + \hat{p}h + \hat{q}h + \check{r}h^2}.$$

Since $\|E^{-1}\| \leq 1/(4\varepsilon + \check{r}h^2)$, we have

$$\|(D - L)^{-1}\| \leq \|E^{-1}\|[1 - \|E^{-1}(L + B)\|]^{-1} \leq \frac{1}{\check{r}h^2} \cdot \frac{4\varepsilon + \hat{p}h + \hat{q}h + \check{r}h^2}{4\varepsilon + \check{r}h^2}.$$

The second factor is a decreasing function of ε, so evaluating the ratio at $\varepsilon = 0$, we obtain

$$\|(D - L)^{-1}\| \leq \frac{\hat{p} + \hat{q} + \check{r}h}{\check{r}^2 h^3}.$$

Each row of U has at most three nonzero entries, so $\|U\| \leq 3\varepsilon$. Combining these inequalities and noting that $\check{r}h \leq \check{r}$, we obtain (4.5).

From the theorem we conclude that for mesh spacings h which satisfy $\varepsilon \leq bh^3/2$, where $b < \check{r}^2/(\hat{p} + \hat{q} + \check{r})$, the error reduction per iteration as measured in the ∞ norm is $< 1/2$. We note that (4.5) is less favorable than (4.3), in that the power of h in the denominator of (4.5) is larger. The reason for the better inequality (4.3) is that the method uses line Gauss Seidel iterations, and the analysis uses the positivity of $\check{p}$.

To use this result, we must develop a method to find an admissible partition of the mesh points. To be practical, the partition should also be of such a character that the corresponding matrices D_i are either of small size or of a type (such as tridiagonal) that the equations (4.4) can be easily solved. The problem of constructing admissible partitions may be put in terms of directed graphs. For this, we extend the notion of α link and β link from pairs of points to pairs of subsets of the set $\mathcal{N}$ of mesh points. If $\mathcal{N}'$ and $\mathcal{N}''$ are disjoint subsets of $\mathcal{N}$, we will say that $(\mathcal{N}',\mathcal{N}'')$ is a β link provided the following condition is satisfied: for $P' \in \mathcal{N}'$ and $P'' \in \mathcal{N}''$, if P' and P'' are neighbors, then (P', P'') is a β link. We say that $(\mathcal{N}'.\mathcal{N}'')$ is an α link if it is not a β link. Thus, $(\mathcal{N}',\mathcal{N}'')$ is an α link if there are a pair of neighboring points $P' \in \mathcal{N}'$ and $P'' \in \mathcal{N}''$ such that the flow vector at P' points away from P''. To any partition $\mathcal{P}$ of $\mathcal{N}$ we associate a directed graph $\mathcal{G}(\mathcal{P})$ as follows. The nodes of $\mathcal{G}(\mathcal{P})$ are the sets of the partition; $(\mathcal{N}',\mathcal{N}'')$ is a directed link of $\mathcal{G}(\mathcal{P})$ if $(\mathcal{N}',\mathcal{N}'')$ is an α link. Then we have

Lemma 2. *A partition of $\mathcal{N}$ can be ordered to be an admissible partition if and only if the graph $\mathcal{G}(\mathcal{P})$ is acyclic.*

Proof. Suppose $\mathcal{G}(\mathcal{P})$ is acycilc; that is, $\mathcal{G}(\mathcal{P})$ has no closed paths. Then there is a set $\mathcal{N}'$ in the partition such that for each $\mathcal{N}''$, $(\mathcal{N}',\mathcal{N}'')$ is not an α link. For otherwise, given an $\mathcal{N}^1$, we could find an $\mathcal{N}^2$ such that $(\mathcal{N}^1,\mathcal{N}^2)$ is an α link, and then we could find an $\mathcal{N}^3$ such that $(\mathcal{N}^2,\mathcal{N}^3)$ is an α link, and continuing in this way, we would create a path that comes back to a previously selected node of $\mathcal{G}(\mathcal{P})$. Let $\mathcal{N}_1 = \mathcal{N}'$ be the first set in the ordering of the partition. The subgraph of $\mathcal{G}(\mathcal{P})$ with $\mathcal{N}_1$ removed also has no closed paths, so there is a set $\mathcal{N}_2$ such that for each $\mathcal{N}'' \neq \mathcal{N}_1, (\mathcal{N}_2,\mathcal{N}'')$ is not an α link. We continue in this way to order the sets of the partition. Let $\mathcal{N}_1,\mathcal{N}_2,\cdots,\mathcal{N}_m$ be the ordering that has been produced. If $k < l$, then $(\mathcal{N}_k,\mathcal{N}_1)$ is not an α link, so $(\mathcal{N}_k,\mathcal{N}_l)$ is a β link, and hence the ordering is admissible. Conversely, if $\mathcal{N}_1,\mathcal{N}_2,\cdots,\mathcal{N}_m$ is an admissible partition, then a closed path in $\mathcal{G}(\mathcal{P})$ must include a pair $(\mathcal{N}_k,\mathcal{N}_l)$ with $k < 1$, and this pair is not an α link, which is a contradiction.

Since the partitions are to be used in s block Gauss Seidel method, it is important to have partitions with samll block size. The following theorem describes some admissible partitions in the case of a general flow vector (p, q). The partitions are "minimal" in the sense that the blocks are as small as possible.

Theorem 3. *Let $\mathcal{P}_m$ be the partition of $\mathcal{N}$ defined by the strongly connected components of $\mathcal{G}_0$. Then $\mathcal{G}_m = \mathcal{G}(\mathcal{P}_m)$ is acyclic, so $\mathcal{P}_m$ is admissible . Furthermore, $\mathcal{P}_m$ is a minimal admissible partition, in the sense that if $\mathcal{Q}$ is another admissible partition of $\mathcal{N}$, then for each set $\mathcal{N}'$ of the partition $\mathcal{P}_m$ there is a set $\mathcal{M}$ of the partition $\mathcal{Q}$ such that $\mathcal{N}' \subset \mathcal{M}$.*

Proof. Suppose $\mathcal{P}_m$ is not admissible. Then $\mathcal{G}_m$ is not acyclic, so there is a cycle of sets of $\mathcal{P}_m$, $\mathcal{N}_i,\cdots,\mathcal{N}_k,\mathcal{N}_{k+1} = \mathcal{N}_1$, such that $(\mathcal{N}_i,\mathcal{N}_{i+1})$ is a link of $\mathcal{G}_m$ for $i = 1,\cdots,k$. Hence there are points $P_i \in \mathcal{N}_i, i = 1,\cdots,k+1$ such that (P_i, P_{i+1}) is a link of $\mathcal{G}_0, i = 1,\cdots,k$. Since each $\mathcal{N}_i$ is a strongly connected component of $\mathcal{G}_m$, each point $Q \in \mathcal{N}_i$ can be connected to P_i through a path of points in $\mathcal{N}_i$, and also P_i can be connected to each $Q \in \mathcal{N}_i$ through a path of points in $\mathcal{N}_i$. Hence any two points in $\mathcal{N}' = \cup\mathcal{N}_i$ can be connected by a path of points in $\mathcal{N}'$, so $\mathcal{N}'$ is a strongly connected set in $\mathcal{G}_0$. This contradicts the assumption that the $\mathcal{N}_i$ are the strongly connected components of $\mathcal{G}_0$. To prove the second assertion, suppose that $\mathcal{Q}$ is an admissible partition, and suppose there is a set $\mathcal{N}'$ of $\mathcal{P}_m$ and two sets $\mathcal{M}', \mathcal{M}''$ of $\mathcal{Q}$ such that $\mathcal{M} \cap \mathcal{N}' \neq \emptyset . \mathcal{M} \cap \mathcal{N}'' \neq \emptyset$. Pick $P' \in \mathcal{M} \cap \mathcal{N}', P'' \in \mathcal{M} \cap \mathcal{N}''$. Since $\mathcal{N}'$ is strongly connected, there is a sequence of points in $\mathcal{N}', P_1 = P', P_2, \cdots, P_k, P_{k+1} = P''$ such that $(P_i.P_{i+1})$ is a link of $\mathcal{G}_0$. Each P_i belongs to a unique set $\mathcal{M}_i$ of the partition $\mathcal{Q}$, so the sequence $\mathcal{M}_1 = \mathcal{M}'$, $\mathcal{M}_2, \cdots, \mathcal{M}_k$, $\mathcal{M}_{k+1} = \mathcal{M}''$ is a path in $\mathcal{G}(\mathcal{Q})$ from $\mathcal{M}'$ to $\mathcal{M}''$. Similarly, there is a path in $\mathcal{N}'$ going from P'' to P', so there is a path in $\mathcal{G}(\mathcal{Q})$ going from $\mathcal{M}''$ to $\mathcal{M}'$. Hence there is a cycle in $\mathcal{G}(\mathcal{Q})$, contradicting the assumption that $\mathcal{Q}$ is an admissible partition.

An algorithm for finding the strongly connected components of a digraph $\mathcal{G}$ is given in [1]. We briefly describe the procedure. The algorithm uses a "depth first search" of the graph $\mathcal{G}$. A depth first search involves selecting a node P of $\mathcal{G}$, a link of $\mathcal{G}$ leading away from P, moving along this link to another node, and so on until one can go no further. (Returning to p, or to another node in the graph that has already been reached, is not allowed.) Then one backs up one step from the last node reached, starts on another link, if there is one, and so on. The purpose of the depth first search is to determine, for each node of $\mathcal{G}$, how many progeny the node has. When the depth first search is completed, the "reverse" digraph $\mathcal{G}_r$ is considered, where $\mathcal{G}_r$ is defined by reversing all the arrows in $\mathcal{G}$. A depth first search is executed on $\mathcal{G}_r$, but this search is started with a node containing the largest number of progeny among the nodes of $\mathcal{G}$. Whenever this depth first search can go no farther without arbitrarily selecting a new node, a strongly connected component of $\mathcal{G}$ has been obtained. The depth first search is then resumed at one of the remaining nodes with a largest number of progeny. The links of the acyclic graph $\mathcal{G}(\mathcal{P}_m)$ are then formed, and the vertices of the graph are arranged in an admissible order. The iteration matrix of the resulting block Gauss Seidel method satisfies (3.2). We have programmed this algorithm and used it to find the minimal admissible partitions $\mathcal{P}_m$ for a variety of flow fields. Some results are given in §6.

§5. Other Flow Directed Schemes

By letting the flow vector guide the ordering of the unknowns, we are led to other block Gauss Seidel methods that are easier to implement than the methods described in §4. While these other orderings are not admissible in the sense of §4, the methods have favorable convergence properties. In this section, we describe 3 iterative methods, denoted FDPI, FDHI, and FDHVI. In the case of FDPI, we give a bound for $\|N\|$ in the one dimensional case. For the other two methods, good estimates for $\|N\|$ remain a problem. Numerical results are given for all the methods in §6.

To describe FDPI, Flow Directed Point Iterations, we divide the set $\mathcal{N}$ of mesh points into 4 subsets as follows:

$$\mathcal{N}_{NE} = \{(i,j) : p_{ij} \geq 0, q_{ij} \geq 0\}; \quad \mathcal{N}_{NW} = \{(i,j) : p_{ij} < 0, q_{ij} \geq 0\};$$
$$\mathcal{N}_{SE} = \{(i,j) : p_{ij} \geq 0, q_{ij} < 0\}; \quad \mathcal{N}_{SW} = \{(i,j) : p_{ij} < 0, q_{ij} < 0\}.$$

We use a Gauss Seidel iteration, ordering the unknowns in the following way. We first sweep through the unknowns in $\mathcal{N}_{NE}$, from left to right and from bottom to top. Next, we sweep through the unknowns in $\mathcal{N}_{NW}$, from right to left and from bottom to top. In the third step we sweep through the nuknowns in $\mathcal{N}_{SE}$, from left to right and from top to bottom. Finally, we sweep through the unknowns in $\mathcal{N}_{SW}$ from right to left and from top to bottom. These 4 steps constitute an iteration of FDPI. In the one dimensional case, FDPI is analysed in Theorem 1 of §3.

The iterative scheme FDHI (Flow Directed Horizontal Iterations) is a variant of line Gauss Seidel. Let $\mathcal{N}_i$ denote the mesh points on the vertical line $x = ih$. We divide $\mathcal{N}_i$ into two subsets:

$$\mathcal{N}_{iE} = \{(i,j) : p_{ij} \geq 0\}; \quad \mathcal{N}_{iW} = \{(i,j) : p_{ij} < 0\}.$$

The FDHI partitioning and ordering of the unknowns consists of the subsets $\mathcal{N}_{iE}$, arranged in order of increasing i, followed by the subsets $\mathcal{N}_{iW}$, arranged in order of decreasing i. Since the difference equations on each of the subsets $\mathcal{N}_{iE}$ or $\mathcal{N}_{iW}$ are a collection of tridiagonal systems, and FDHI iteration is not expensive. By considering the mesh points $\mathcal{N}_j$ on a horizontal line $y = jh$ and dividing $\mathcal{N}_j$ into subsets $\mathcal{N}_{JN}, \mathcal{N}_{jS}$, we may construct a similar iterative scheme FDVI. We designate by FDHVI the iterative scheme consisting of alternate iterations of FDHI and FDVI. Note that in the course of one iteration of FDHVI, the unknown at each mesh point has been updated twice; we have taken this into account in our evaluation of the method.

It would seem that there is a close relation between the methods described above and the symmetric Gauss Seidel method. Let us denote by SHI the symmetrized version of line Gauss Seidel, where the lines are vertical mesh lines, and the iterations are first, left to right, and second, right to left. The principal difference between FDHI and SHI is that in FDHI, on the rightward sweep, only those unknowns for which the flow vector points to the right are updated, whereas, with SHI, every unknown is updated on each sweep. The numerical results in §6 include some results with SHI.

§6. Numerical Results

To illustrate some features of the methods discussed above, we consider four problems. Since the analysis suggests that flow directed iterations are at their worst when used with flow fields containing sources, the flow fields of the four problems will be chosen to exhibit various types of sources. The four flow fields are given on the unit square by the following formulas. In each case we take $r(x,y) = .5$.

Problem 1. $p(x,y) = 3x - y - 1, q(x,y) = 1$.

Problem 2. $p(x,y) = 3x - y - 1, q(x,y) = -x - 3y + 2$.

Problem 3. $p(x,y) = 2(x - .5) - \rho(y - .5), q(x,y) = \rho(x - .5) + 2(y - .5), \rho = [(x - .5)^2 + (y - .5)^2]^{1/2}$.

Problem 4. $p(x,y) = -2(x - .5) - \rho(y - .5)$, $q(x,y) = \rho(x - .5) - 2(y - .5)$, $\rho = [(x - .5)^2 + (y - .5)^2]^{1/2}$.

The stream lines of the flow fields in Problems 3 and 4 are expanding spirals $\rho = 2\theta$ (for Problem 3) and contracting spirals $\rho = -2\theta$ (for Problems 4), where (ρ, θ) are polar coordinates centered at the point $(.5, .5)$. Schematic illustrations of the four flow fields are given in Figures 1a–4a.

The first set of results are not numerical. Rather, they show the minimal admissible partition $\mathcal{P}_m$ on a 5×5 mesh, in each of the four cases. We display, in Figures 1b–4b, the 5×5 mesh, with the (schematic) direction of the flow vector. We also show the partition of the mesh points into the subsets of $\mathcal{P}_m$. In Figures 1c–4c, we show the nodes and links of the reverse digraph $\mathcal{G}_r(\mathcal{P}_m)$. In perusing these figures, notice how the ordering of the partitions follows the direction of the flow, and notice also how aggregating the mesh points into the sets of $\mathcal{P}_m$ handles the problem of flows emanating from groups of mesh points. In Problem 3, the source of flow at the center of the expanding spiral requires a large block of points (8 points) in the second set of $\mathcal{P}_m$. There is no guarantee that the block size remains bounded as the mesh spacing becomes finer. The flow field of Problems 4, which represents a contracting spiral, does not have this problem of large block size.

The next set of results is contained in Table 1–4. They show the error in the solution, for Problems 1–4, on a 20×20 mesh, after 10 iterations, for various values of ε. The methods FDPI, FDHI, FDHVI, and SHI are discussed in §5. We have also included two more methods, denoted HI and MPPI. By HI (Horizontal Iterations) we mean the line Gauss Seidel method where the line are the vertical lines $j =$ constant, and where the lines are swept in the usual left-to-right order. By MPPI (Minimal Partition Point Iterations) we mean the point Gauss Seidel method in which the points are ordered in the manner given by our program that provides the minimal partition $\mathcal{P}_m$. Note that the analysis of the Gauss Seidel method arising from $\mathcal{P}_m$ says nothing about what would happen if the ordering given by $\mathcal{P}_m$ is used to generate a point Gauss Seidel method. Nevertheless, the numerical results show that MPPI performs quite well.

The concept of what constitutes an "iteration" needs to be explained to understand the tables. We have counted as one "iteration" a sweep of the mesh in which each unknown is changed once. Thus a complete step of FDHVI counts as two iterations. The same holds for a complete step of SHI. It should be emphasized that the various iterations may not take the same amount of work per iteration. The way FDHI has been programmed, for example, requires a calculation for each mesh point and an update only at the mesh points for which the flow vector points in the appropriate direction. We may note that HI performs very poorly, especially for decreasing ε. The other methods perform well, and FDHVI seems to perform the best. In Problem 4, the blocks of the minimal admissible partition $\mathcal{P}_m$ are all of size 1. This means that for this problem, the method MPPI satisfies the hypotheses of Theorem 2. In the Tables we see that MPPI performs very well for Problem 4. Note, however, that MPPI performs well for the other problems, even though the hypotheses of Theorem 2 are not satisfied. It seems that the ordering provided by the minimal admissible partition is a good one to use for point Gauss Seidel iterations for a variety of flow fields.

Problem 1. Error After 10 Iterations

$P(x,y) = 3x - y - 1, \quad q(x,y) = 1, \quad r(x,y) = .5$						
eps	HI	FDPI	FDHI	FDHVI	SHI	MPPI
.1	7.55E–02	9.34E–01	8.25E–01	7.80E–01	7.96E–01	9.32E–01
.1–1	1.60E–01	2.24E–01	2.40E–02	2.17E–02	8.39E–02	2.25E–01
.1–2	1.23E–01	3.33E–06	1.57E–07	3.27E–08	1.02E–05	3.85E–06
.1–3	7.64E–02	1.00E–10	1.29E–09	1.16E–13	1.97E–09	5.11E–10
.1–4	7.05E–02	2.34E–12	5.97E–10	1.03E–18	5.02E–10	4.46E–11
.1–5	6.97E–02	1.25E–12	5.50E–10	1.01E–23	4.37E–10	3.11E–11

Problem 2. Error After 10 Iterations

$P(x,y) = 3x - y - 1, \quad q(x,y) = -x - 3y + 2, \quad r(x,y) = .5$						
eps	HI	FDPI	FDHI	FDHVI	SHI	MPPI
0.1	7.10E–02	9.03E–01	7.75E–01	6.85E–01	7.25E–01	9.02E–01
0.1–1	9.36E–02	2.44E–01	6.10E–02	3.52E–02	1.09E–01	2.38E–01
0.1–2	2.24E–01	4.10E–04	6.78E–05	1.52E–06	1.07E–03	3.57E–04
0.1–3	5.41E–01	1.71E–08	1.18E–08	2.06E–12	1.31E–05	1.59E–08
0.1–4	6.26E–01	1.31E–09	4.87E–10	3.34E–17	1.07E–05	3.69E–10
0.1–5	6.35E–01	9.96E–10	4.40E–10	2.23E–21	1.05E–05	2.59E–10

Problem 3. Error After 10 Iterations

Spiral with $a = 2$, center $= (.5, .5), r(x,y) = .5$						
eps	HI	FDPI	FDHI	FDHVI	SHI	MPPI
0.1	2.60E–02	9.53E–01	8.87E–01	8.93E–01	8.49E–01	9.52E–01
0.1–1	5.17E–02	6.44E–01	4.20E–01	4.48E–01	4.57E–01	6.50E–01
0.1–2	1.96E–01	5.45E–02	6.39E–03	9.47E–03	1.99E–02	5.91E–02
0.1–3	4.56E–01	4.50E–06	6.49E–08	1.83E–07	2.37E–05	6.54E–06
0.1–4	5.25E–01	1.91E–11	1.09E–12	5.63E–13	8.26E–07	9.69E–10
0.1–5	5.32E–01	1.49E–13	1.06E–13	1.52E–17	5.92E–07	1.04E–10

Problem 4. Error After 10 Iterations

Spiral with $a = -2$, center $= (.5,.5), r(x,y) = .5$						
eps	HI	FDPI	FDHI	FDHVI	SHI	MPPI
0.1	8.37E–02	8.45E–01	5.55E–01	5.66E–01	6.61E–01	8.48E–01
0.1–1	1.65E–01	1.42E–01	9.94E–03	1.24E–02	4.61E–02	1.46E–01
0.1–2	4.65E–01	2.35E–06	4.10E–09	3.30E–09	1.48E–05	2.50E–06
0.1–3	6.14E–01	2.79E–15	2.19E–18	7.09E–19	3.88E–10	2.06E–15
0.1–4	6.33E–01	9.10E–25	6.10E–28	4.86E–28	4.44E–15	2.18E–25
0.1–5	6.35E–01	7.53E–34	4.41E–37	2.84E–35	4.51E–20	2.18E–35

Acknowledgement of Support. The work of H. Han was supported by the National Natural Science Foundation of China and by the U.S. National Science Foundation, grant INT-8517582. The work R.B. Kellogg was supported in part by the U.S. National Science Foundation, Grant INT-8517582.

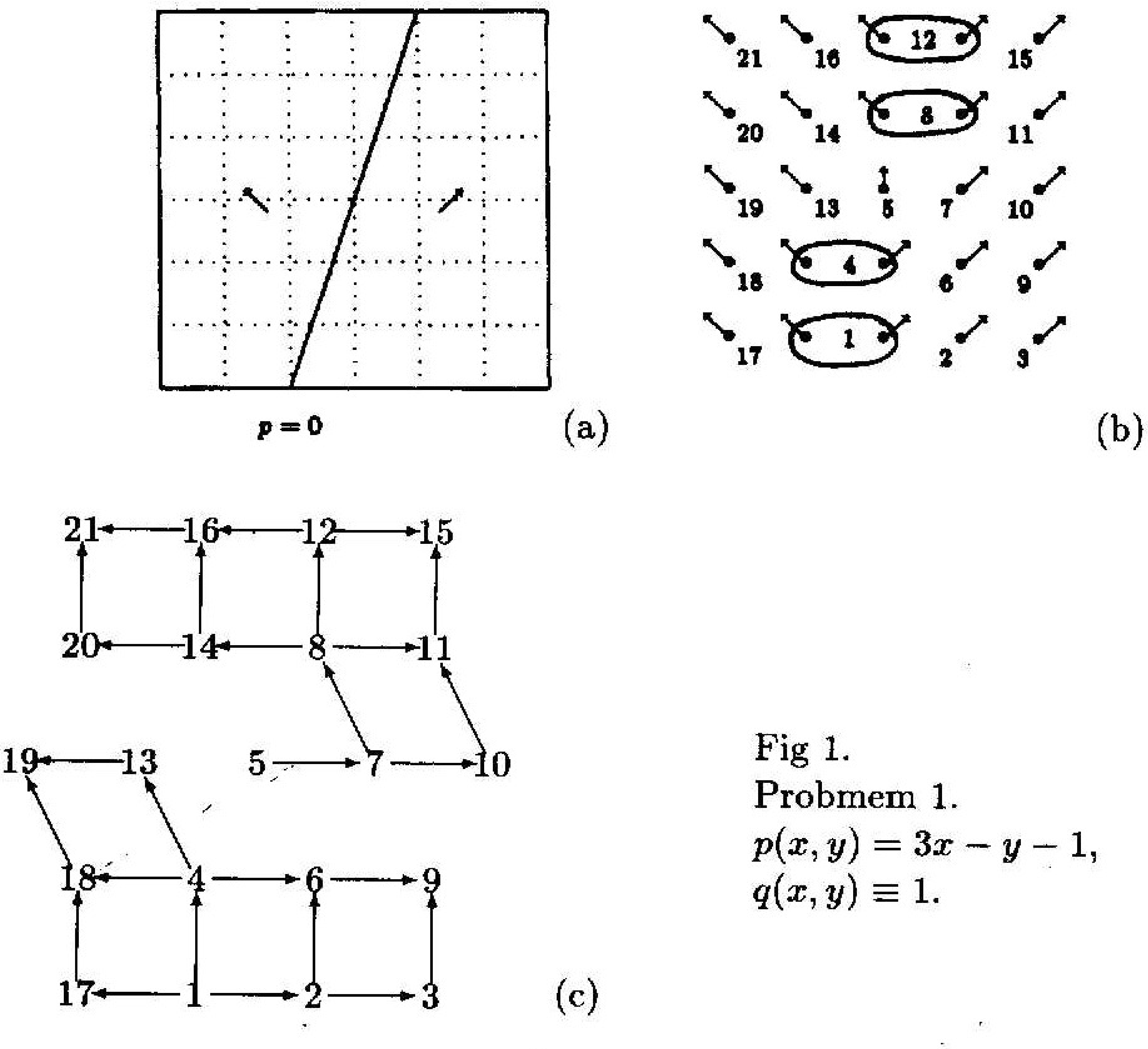

Fig 1.
Probmem 1.
$p(x,y) = 3x - y - 1,$
$q(x,y) \equiv 1.$

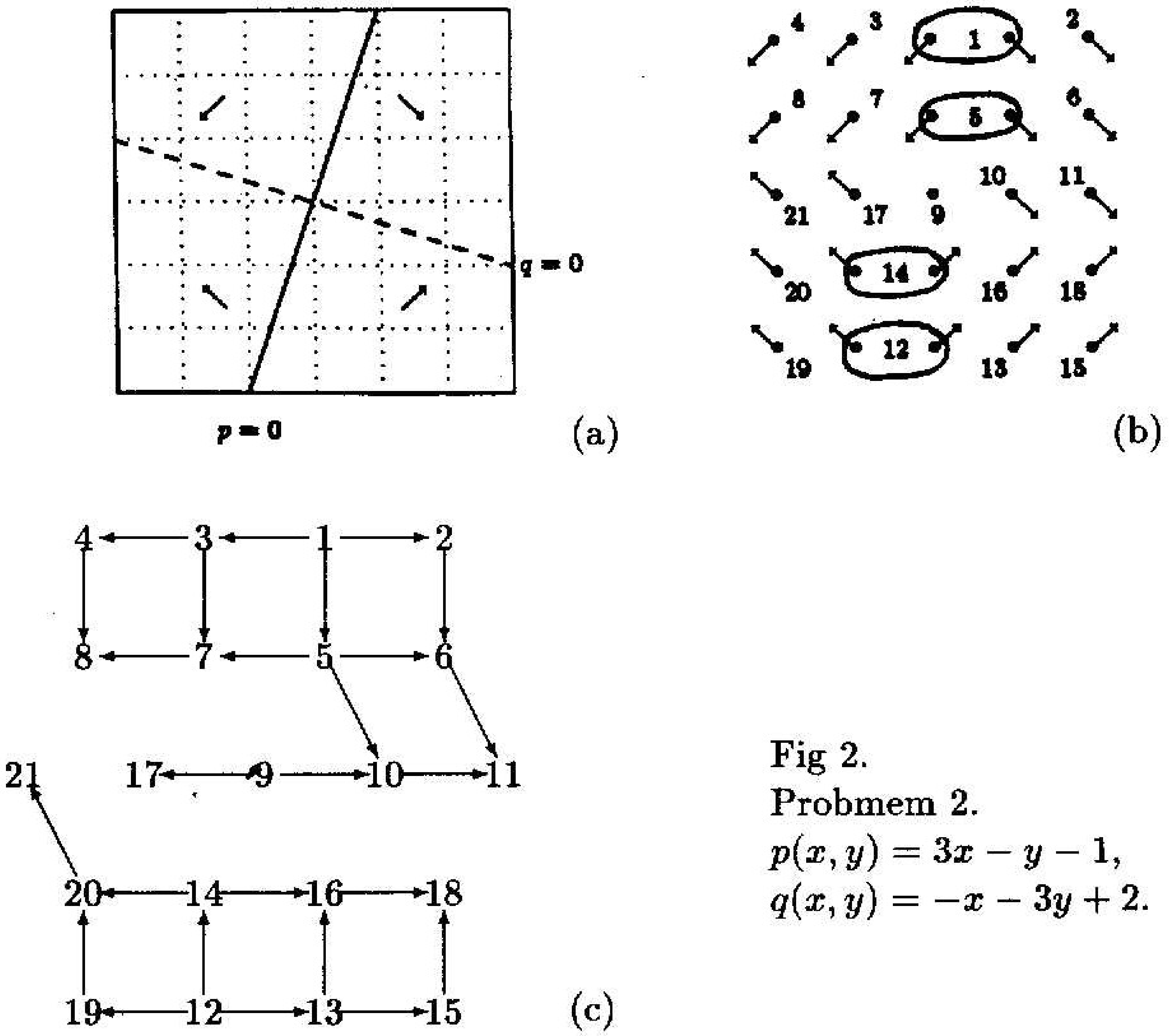

Fig 2.
Probmem 2.
$p(x, y) = 3x - y - 1,$
$q(x, y) = -x - 3y + 2.$

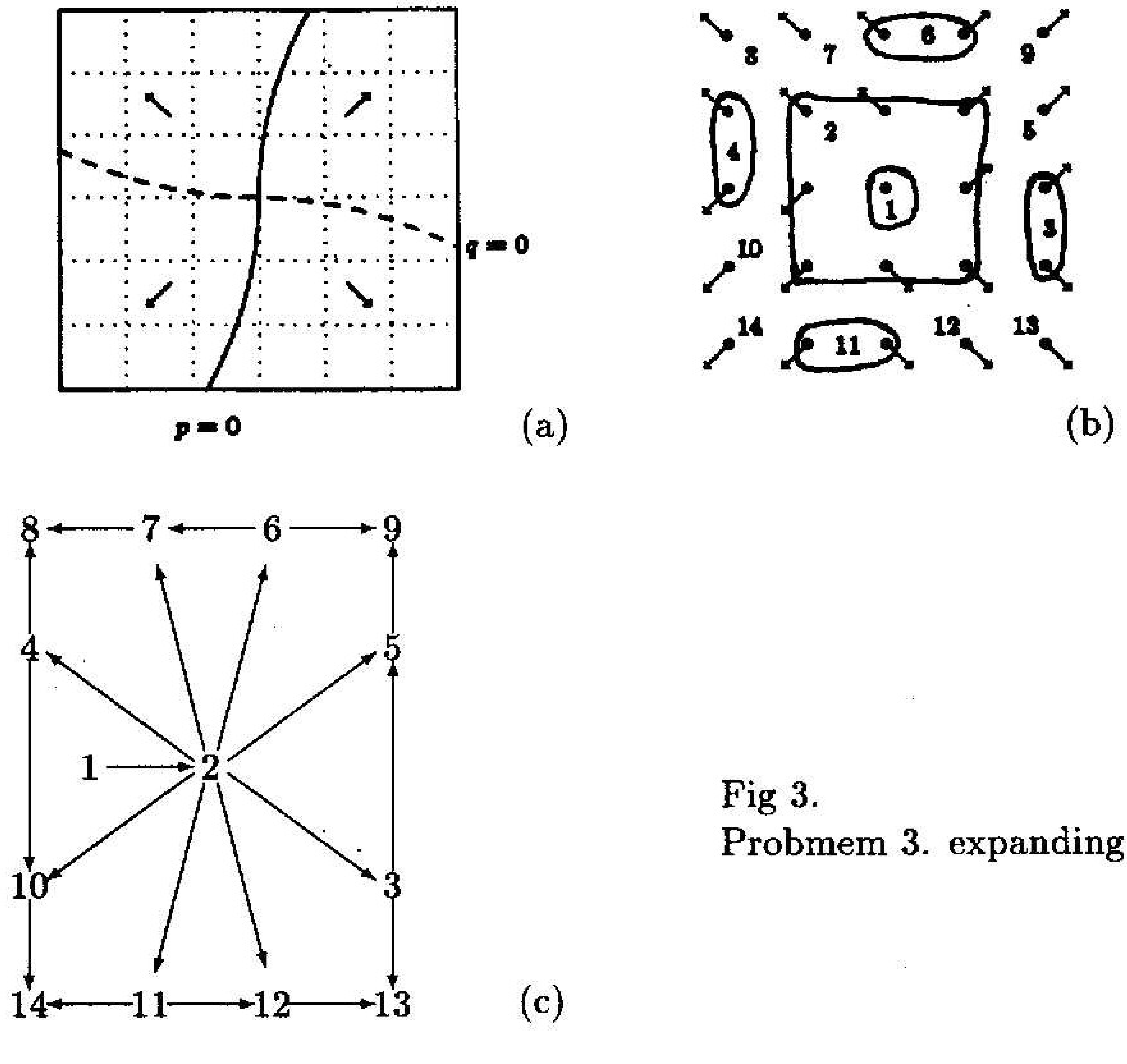

Fig 3.
Probmem 3. expanding spiral

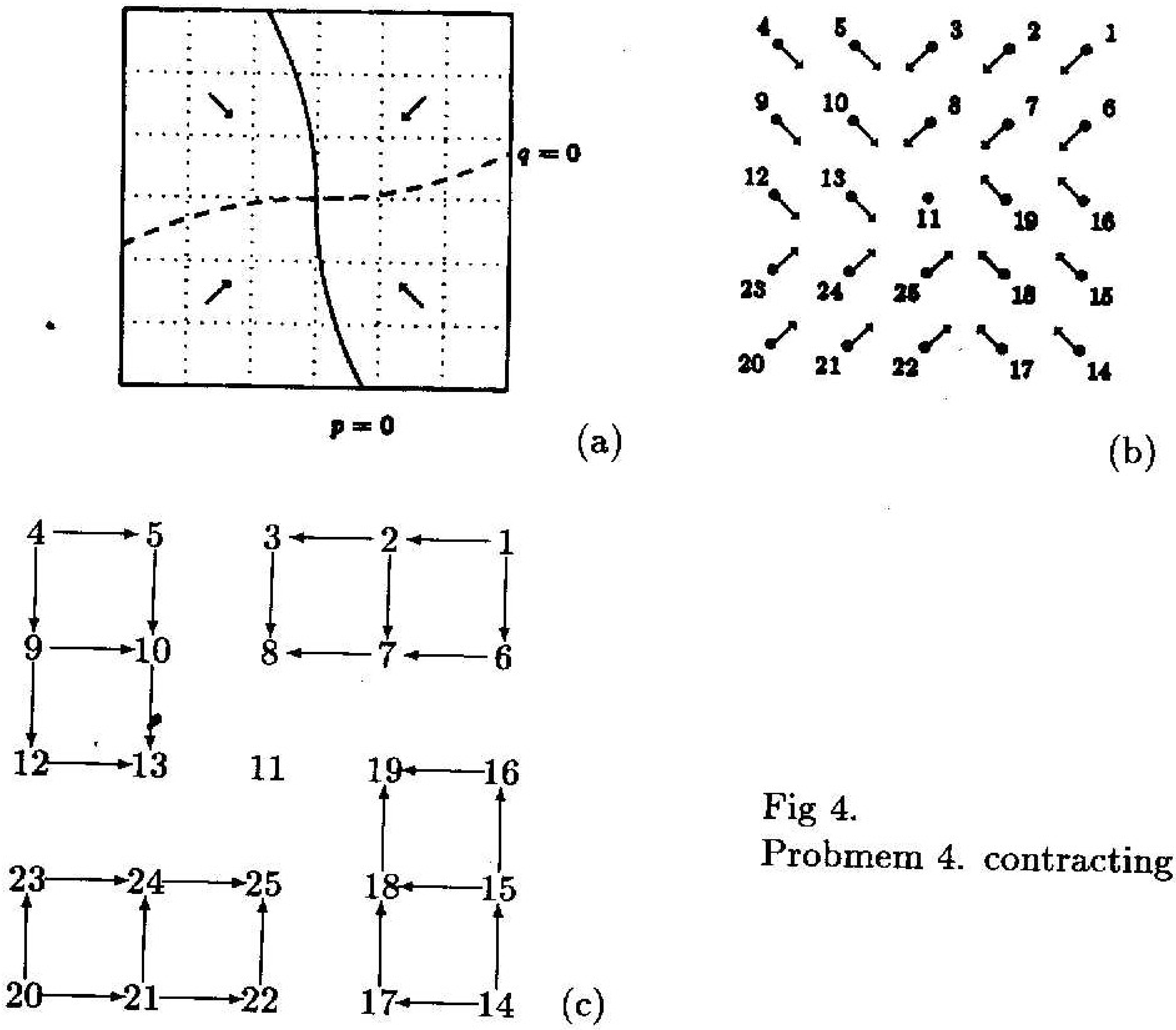

Fig 4.
Probmem 4. contracting spiral

References

[1] A.V. Aho, J.E. Hopcroft and J.D. Ullman, The design and analysis of computer algorithms, Addison Wesley, Reading, Mass., 1974.

[2] A.E. Berger, J.M. Solomon and M. Ciment, Uniformly accurate difference methods for a singular perturbation problem, in Boundary and Interior Layers, J.J. H.Millar, Editor, Boole Press, Dublin Ireland, 1980, 14–28.

[3] A.E. Berger, H.Han and R.B. Kellogg, A priori estimates and analysis of a numerical method for a turning point problem, *Math Comp.*, **42** (1984), 465–492.

[4] V.Ervin and W. Layton, On the approximation of derivatives of singularly perturbed boundary value problems, *SIAM J. Sci. Stat. Comput.*, **8** (1987), 265–277.

[5] V. Ervin and W. Layton, A second order accurate, positive scheme for singular perturbed boundary value problems, Technical Report, School of Mathematics, Georgia Tech., Atlanta GA, 1985.

[6] C.I. Goldstein, Preconditioning elliptic operators with dominant low order terms, Advances in Computer Methods for Partial Differential Equations VI, V. R. Vichnevetsky and R.S. Stepleman, Eds., IMACS, 1987, 67–69.

[7] R.B. Kellogg and A. Tsan, Analysis of Some difference approximations for a singular perturbation problem without turning points, *Math Comp.*, **32** (1978), 1025–1059.

[8] John Strikwerda, Iterative methods for the numerical solution of second order elliptic equations with large first order terms, *SIAM J. Sci. Stat. Comp.*, **1** (1980), 119–130.

[9] R.S. Varga, Matrix Iterative Analysis, Prentice Hall, Englewood Cliffs, NJ, 1962.

[10] Wen-Ting Tong, On the spectral radius of matrices, *Linear and Multilinear Algebra,* **20** (1987), 175–182.