Orderings for
Conjugate Gradient Preconditionings

James M. Ortega

Computer Science Report No. TR-90-24
August 14, 1990

Research supported in part by the National Aeronautics and Space Administration under NASA Grant NAG-1-1112-FDP.

Orderings for Conjugate Gradient Preconditionings

James M. Ortega

1. Introduction

We are concerned in this paper with solution of the large sparse linear system

$$A\,\mathbf{x} = \mathbf{b} \tag{1.1}$$

by the conjugate gradient method with SSOR or incomplete Cholesky (IC) preconditioning. We assume that the system (1.1) arises from the discretization of a Poisson-type equation of the form

$$\nabla(\mathbf{K} \cdot \nabla u) = f \tag{1.2}$$

in two or three dimensions. Here $\mathbf{K}$ is a given vector-valued function of the spatial variables such that (1.2) is elliptic. For simplicity, we will restrict ourselves to rectangular or parallelpiped domains and Dirichlet boundary conditions, although many of the considerations are more general. We also assume that finite difference discretizations are done in such a way that A is symmetric positive definite, and has the usual five-diagonal structure in two dimensions and seven-diagonal structure in three dimensions.

The preconditioning step in the conjugate gradient method requires solving a subsidiary system of equations

$$M\,\tilde{\mathbf{r}} = \mathbf{r} \tag{1.3}$$

to obtain a modified residual vector $\tilde{\mathbf{r}}$. M is the preconditioning matrix, assumed to be symmetric positive definite, and for many commonly used preconditioners, such as SSOR or IC factorization, has the form

$$M = L\,D\,L^T \tag{1.4}$$

where L is lower triangular and D is diagonal. Thus, the solution of (1.3) requires the solution of triangular systems with coefficient matrices L and L^T. In many cases, for example, SSOR and no-fill IC factorization, L will have the same non-zero structure as the lower triangular portion of A itself. The problem, then, is how to solve such triangular systems effectively on parallel and vector architectures.

There have been two main approaches to this problem. The first is to reorder the unknowns so that the coefficient matrix takes the form $\hat{A} = PAP^T$ for some permutation matrix P. The classical such reordering is the red-black ordering (Young [1971]) in which

$$\hat{A} = \begin{bmatrix} D_1 & C^T \\ C & D_2 \end{bmatrix} \tag{1.5}$$

where D_1 and D_2 are diagonal. The use of the red/black ordering for carrying out the SOR iteration on parallel and vector machines dates back to the early 1970's (Erickson [1972], Lambiotte [1975]). More recently, various *multicolor orderings* (see, e.g. Ortega [1988] for a review) have been used for both Poisson-type equations or more general equations with more general finite element or finite difference discretizations. For a multicolor ordering, the coefficient matrix takes the form

$$\hat{A} = \begin{bmatrix} D_1 & C_{21} & \cdot & \cdot & \cdot & C_{c1} \\ C_{21} & D_2 & & & & \\ \cdot & & \cdot & & & \\ \cdot & & & \cdot & & \\ \cdot & & & & \cdot & C_{c,c-1}^T \\ C_{c1} & \cdot & \cdot & \cdot & C_{c,c-1} & D_c \end{bmatrix} \tag{1.6}$$

where, again, the D_i are diagonal, and c is the number of colors. For a multicolor ordering, the

solution of a lower triangular system $L\mathbf{x} = \mathbf{d}$ of the same structure can be computed by

$$\mathbf{x}_i = D_i^{-1}(\mathbf{d}_i - \sum_{j<i} L_{ij}\, \mathbf{x}_j) \ , \quad i = 1, \cdots, c \tag{1.7}$$

Since the D_i are diagonal, the solution has been reduced to matrix-vector multiplys, which are potentially ideal for parallel and vector machines. In particular, for the red-black ordering, (1.7) reduces to

$$\mathbf{x}_2 = D_2^{-1}\ (\mathbf{d}_2 - C\,\mathbf{x}_1) \tag{1.8}$$

For regular problems, C consists of only a few non-zero diagonals and the multiplication $C\mathbf{x}_1$ is efficiently executed by multiplication by diagonals (Madsen et al. [1976]).

Multicolor orderings exhibit a high degree of parallelism but may have a deleterious effect on the rate of convergence of iterative methods. For SOR itself, the asymptotic rate of convergence for the red/black ordering is the same as the natural ordering (Young [1971]) and the same result extends to many multicolor orderings (Adams and Jordan [1985]). Moreover, the use of multicolor orderings seems to enhance the rate of convergence in practice. Unfortunately, there has been growing evidence that the rate of convergence of the conjugate gradient method may be degraded, sometimes seriously, when such orderings are used for preconditioners (Poole and Ortega [1987], Ashcraft and Grimes [1988]). For example, for the equation (1.2) on a $63 \times 63 \times 63$ grid (250,000 unknowns), Harrar and Ortega [1990] reported 162 iterations for SSOR preconditioned conjugate gradient using the red/black ordering and 38 iterations using the natural ordering. Moreover, the red/black ordering is the basis for the reduced system conjugate gradient method (see, e.g., Hageman and Young [1981]), and since this method is mathematically equivalent to SSOR preconditioning on the original system (see, e.g., Harrar and Ortega [1990]), it suffers from the same rate of convergence problem.

A number of other reorderings have been considered but with, mostly, similar results. Duff and Meurant [1989] reported on an extensive set of experiments using incomplete Cholesky preconditioning for the problem (1.2) on a 30×30 grid in two dimensions, with problems containing anisotropy, discontinuous coefficients, etc. They considered 16 different reordering strategies of which only six gave rates of convergence comparable to the natural ordering on all problems, and three of these orderings are equivalent to the natural ordering.

The second main approach to solving the triangular systems effectively is to obtain what parallelism is available in the natural ordering. The basic idea is exemplified by the diagonal ordering shown in Figure 1.1 in two dimensions.

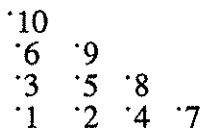


Figure 1.1 Diagonal Ordering of Grid Points

In carrying out SOR with a five point stencil the unknowns on each diagonal can be updated in parallel, or with vector operations whose lengths are the number of points in a diagonal. This is true because the five-point stencil couples unknowns only from different diagonals. However, the updates produced from this ordering are exactly the same as with the natural ordering. For example, the update at point 5 in Figure 1.1 depends only on the updated values at 2 and 3 and the old values at 8 and 9; thus, it makes no difference whether all the unknowns on the first row are updated before 5 is or not. This basic idea extends to three dimensions and is also related to multicolor orderings, as will be discussed in the next section. In two dimensions, this diagonal ordering was studied in Young [1971] as an example of a consistent ordering and discussed as a possible paradigm for vectorization by Hayes [1978]. It

has been explored in increasingly sophisticated ways by van der Vorst [1983], [1989a], [1989b], Schlichting and van der Vorst [1989], and Ashcraft and Grimes [1988], in both two and three dimensions. We will discuss the diagonal ordering in more detail in the next section, including extensions to more general discretizations.

Given the experimental results that certain orderings may degrade the rate of convergence, the question is why this is the case. One attempt at an explanation was Melhem [1986] and the concept of "zero-stretch." He observed that orderings, such as the red/black ordering, that moved elements of the matrix away from the main diagonal tended to degrade the rate of convergence. However, we follow here the lead of Duff and Meurant [1989], who computed the "remainder matrices" for their different orderings. We discuss in Section 4, for a few particular orderings, the structure and size of these remainder matrices. Prior to that, we collect in Section 3 some results on estimation of the condition number as a function of reorderings.

2. Diagonal and Related Orderings

In this section, we expand on the discussion of diagonal orderings, as exemplified by Figure 1.1. As observed by Poole and Ortega [1987], the fact that points on a diagonal are not coupled with themselves implies that the ordering of Figure 1.1 is a multicolor ordering if we assign a separate color to each diagonal. The coefficient matrix in this ordering has the form shown in Figure 2.1, which is taken from Stotland [1990]. Note that this ordering is just the Cuthill-McKee (CM) [1969] ordering for bandwidth/profile minimization. Duff and Meurant [1989] also consider the reverse CM ordering, in which the numbering proceeds from the opposite vertex of the grid, as well as a block CM ordering. These are also equivalent to the natural ordering.

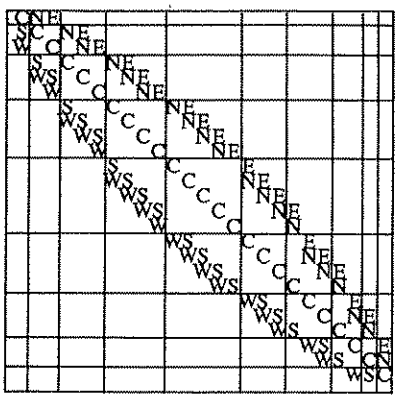


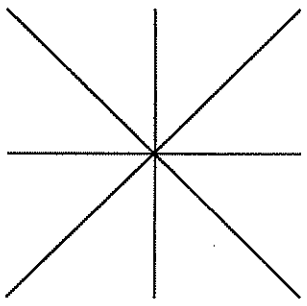
Figure 2.1 Matrix for Diagonal Ordering (Two Dimensions)

Figure 2.1 illustrates the vectors that can be used to carry out the multiplications of (1.7). In particular, the vector lengths are 1, ⋯ , N–1, N, N–1, ⋯ ,1 on an N × N grid. These are considerably smaller than the corresponding vector lengths in the red/black ordering, which are $O(N^2/2)$. But Ashcraft and Grimes [1988] have given results on a CRAY X- MP indicating that the diagonal ordering is superior to the red/black ordering on that machine. However, Elman and Agron [1989] and Chan et al. [1989] have shown that the red/black or multicoloring orderings will probably be superior on highly parallel machines. In particular, the first of these two papers gives theoretical results for a hypercube architecture, and the second gives theoretical and computational results for a Connection Machine. Thus, the situation seems to be that the diagonal ordering will be superior on machines requiring only a modest degree of parallelism or vectorization but as the parallelism of the architecture increases, red/black or multicoloring orderings may become relatively more competitive.

We next discuss a problem with the diagonal ordering and a way to alleviate it that was observed by van der Vorst [1989a]. For the reordered matrix of Figure 2.1, we no longer have the long diagonals of the naturally ordered matrix to use in the matrix-vector multiply in the conjugate gradient method. In particular, there are breaks in the diagonals, corresponding to the

different B_{ij} of (1.6). On the other hand, we can leave the matrix in the naturally ordered form and still carry out the updates according to the diagonal ordering paradigm. This is the approach taken by Ashcraft and Grimes [1988], but the vectors now have stride N−1, which may cause memory bank conflicts on Cray machines. Thus, in the natural ordering, the data is arranged optimally for the matrix multiply but not for the preconditioning and vice-versa for the diagonal ordering. However, as observed by van der Vorst [1989a], we can circumvent this problem by means of the Eisenstat modification [1981], which eliminates the need for the matrix multiply. Thus, we can use the diagonal ordering of Figure 2.1 so as to have suitable vectors for the preconditioning.

Ordering by diagonals is a special case of *wavefront* methods, which can be applied to more general triangular systems; see, Greenbaum [1986] and Saltz [1990]. Rather than the general case, we next discuss another particular discretization. Consider the nine point stencil shown in Figure 2.2, which is used for fourth order approximations to Poisson's equation or for equations containing a mixed derivative u_{xy}. A multicolor "diagonal" ordering for the nine-point stencil is also shown in Figure 2.2.



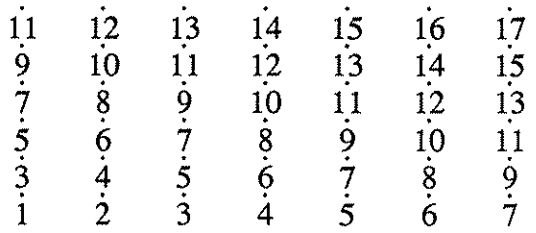


Figure 2.2. Nine-Point Stencil and Multicolor "Diagonal" Ordering

Again for SOR, unknowns corresponding to grid points of the same number in Figure 2.2 can be updated simultaneously. For example, after unknowns 1 and 2 have been updated, both

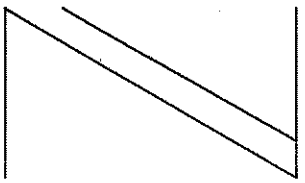
of the unknowns labeled 3 can be updated simultaneously since the only updated unknowns they depend on are 1 and 2. The only difference between this situation and the previous one with the five-point stencil is that now the set of grid points of color i is no longer a diagonal of adjacent grid points but, rather a bent "diagonal" of points separated by a "knight's move", as shown in Figure 2.2.

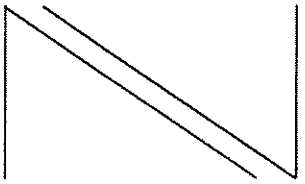
If the unknowns are ordered corresponding to the multicoloring of Figure 2.2, the coefficient matrix is again a multicolored matrix and takes the form

$$A = \begin{bmatrix} D_1 & L_{12} & L_{13} & L_{14} & & & \\ & D_2 & L_{12} & L_{13} & L_{14} & & \\ & & D_3 & L_{12} & L_{13} & L_{14} & \\ & & & D_4 & L_{12} & L_{13} & L_{14} \\ & & & & \ddots & \ddots & \ddots \\ & & & & & & D_c \end{bmatrix} \quad \text{Symm} \tag{2.1}$$

The D_i are diagonal: D_1 and D_2 are 1×1 matrices, D_3 and D_4 are 2×2, and so on to the maximum size, which is O(N/2), after which they decrease in size. More precisely, on an $N \times N$ grid, if N is even, there are $3N - 2$ colors (i.e. $c = 3N - 2$ in (1.6)) and the maximum size of the D_i is N/2, which is taken on by N+2 of the D_i. If N is odd, $c = 3N$ and the maximum size of the D_i is $(N + 1)/2$, which is taken on by (N+3)/2 of the D_i.

The off-diagonal matrices $L_{i,i+2}$ and $L_{i,i+3}$ in (2.1) each have a single non-zero diagonal whose length is the same as the corresponding D_i. Except for the 1×1 matrix L_{12}, the matrices $L_{i,i+1}$ each have two non-zero diagonals and the precise structure depends on whether i is odd or even as shown in (2.2) and (2.3).

L_{ii+1} =  i odd, $\frac{i+1}{2} \times \frac{i+1}{2}$ (2.2)

L_{ii+1} =  i even, $\frac{i}{2} \times \frac{i}{2} + 1$ (2.3)

Thus, $L_{i,i+1}$ has diagonals of length i/2 if i is even, and (i+1)/2 and (i−1)/2 if i is odd; this holds until the $L_{i,i+1}$ reach their maximum size and then start decreasing. Note that $L_{i,i+1}$ is square if i is odd but rectangular if i is even, as shown by the matrix sizes given in (2.2) and (2.3). This is illustrated in Figure 2.3, which shows the first few blocks of a typical matrix.

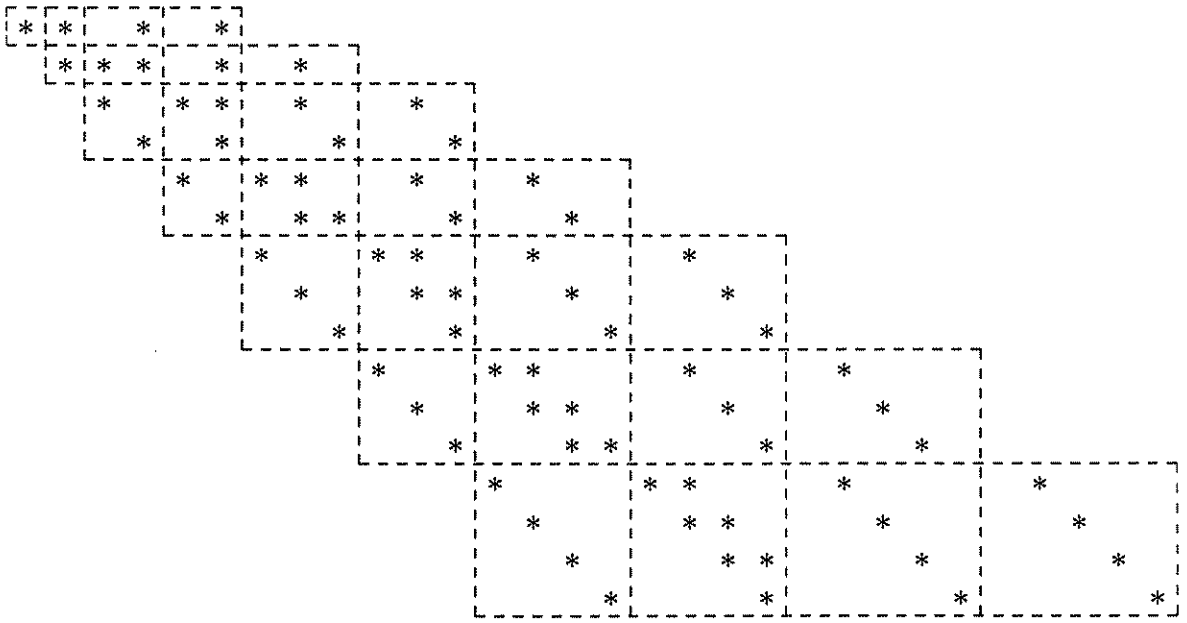


Figure 2.3 Structure of Matrix

We next comment on the differences in vector lengths between the five-point and nine-point stencils. With the nine-point stencil we have noted that the maximum size of a D_i is

O(N/2); this is the maximum vector length if we don't attempt to couple diagonals between adjacent blocks. This contrasts to the maximum vector length of O(N) for the five-point stencil. Thus, the use of the nine-point stencil approximately halves the vector lengths, as compared with the five-point stencil. This is analogous to the use of the red/black ordering for the five-point stencil which gives vector lengths of $O(N^2/2)$ compared with a four-color ordering for the nine-point stencil, which gives vector lengths of $O(N^2/4)$, nominally half the vector length of the red/black ordering.

We next consider three-dimensional problems. For simplicity we will restrict ourselves to Poisson-type equations (1.1) and the seven-point stencil. This stencil is the natural extension to three-dimensions of the five-point stencil for two dimensions.

Ordering the grid points by diagonals in two dimensions extends in a natural way for three dimensions to ordering by diagonal planes. (See Ashcraft and Grimes [1988] and van der Vorst [1989a].) This is illustrated in Figure 2.4 for a $3 \times 3 \times 3$ grid. Only two diagonal planes are shown in Figure 2.4 but there are 7 such planes, including two which consist of only a single point. The grid points in these seven planes are shown in Figure 2.5. The first three planes in Figure 2.5 correspond to the planes shown in Figure 2.4, including the first plane which contains only one point.

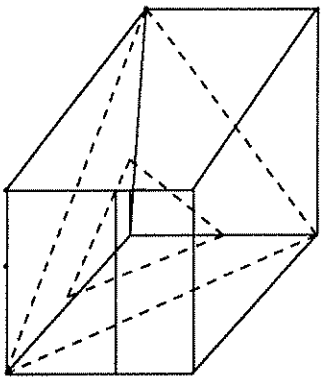


Figure 2.4 Ordering by Diagonal Planes; $3 \times 3 \times 3$ grid

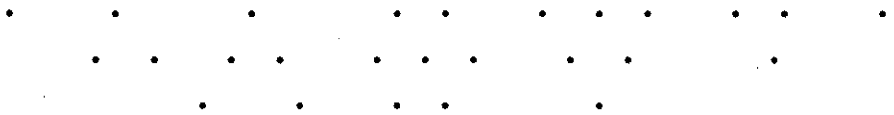


Figure 2.5 Grid Points in Diagonal Planes: $3 \times 3 \times 3$ grid

Points in a plane are assigned the same color. For the $3 \times 3 \times 3$ example, there are 7 planes and, hence, 7 colors. For an $N \times N \times N$ grid there are $3N-2$ planes and $3N-2$ colors. If we order the grid points, and therefore the unknowns, corresponding to these diagonal planes, we again obtain a multicolor block tridiagonal matrix. The detailed structure of the matrix in the case of a $4 \times 4 \times 4$ grid is shown in Figure 2.6, which is taken from Stotland [1990].

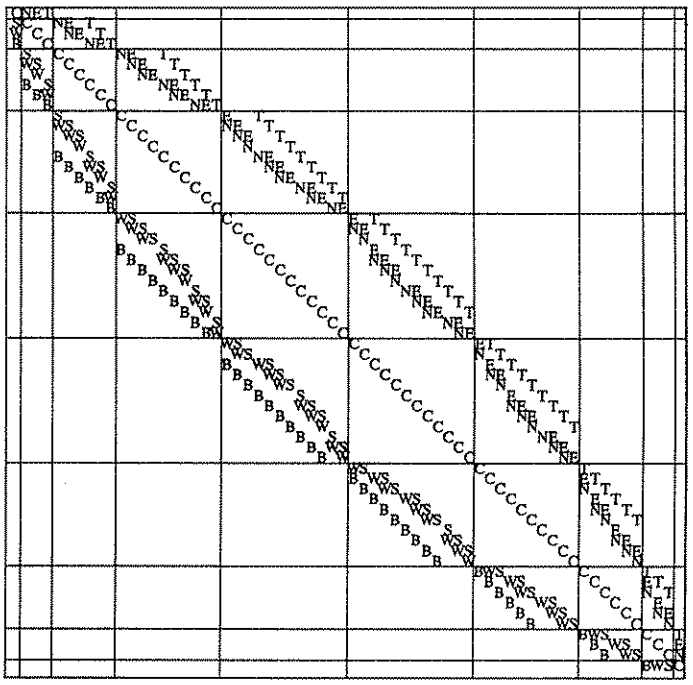


Figure 2.6 Matrix for Diagonal Ordering - Three Dimensions - 4×4 Grid

The maximum number of grid points in a plane is $3N^2/4$ if N is even and $(3N^2+1)/4$ if N is odd. Since all points in a plane can be updated by SOR in parallel, the maximum degree of parallelism is $O(3N^2/4)$. However, an examination of Figure 2.6 shows that the same is not true

of the vector lengths. Consider one of the off-diagonal blocks of maximum size. Although there is one long diagonal, the other diagonals break up into lengths that are associated with two-dimensional problems. There is some over-lapping of the diagonals across the submatrices corresponding to the two dimensional problems so that some vectors of length greater than N can be obtained. Moreover, the gather operation can be used on some of the shorter vectors. However, in three dimensions the vectorization properties of the diagonal ordering are not as good as the parallel properties. This is in contrast to the two dimensional situation.

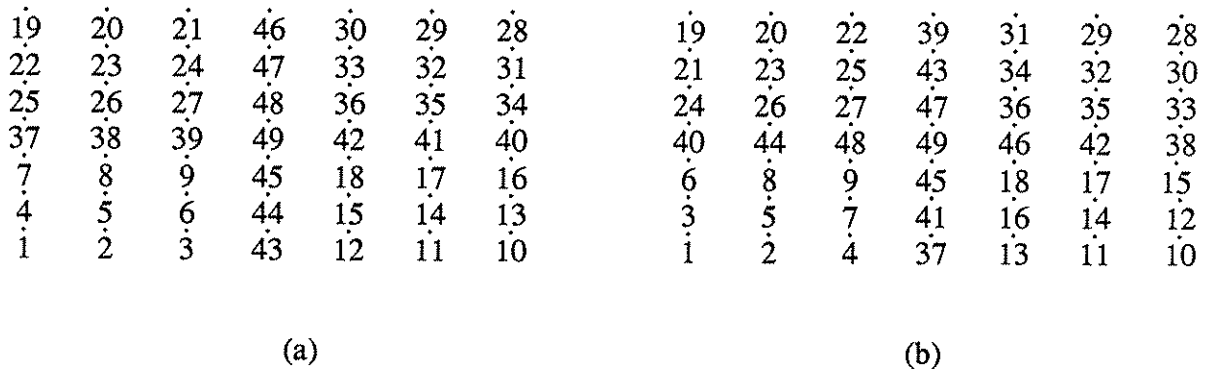


Figure 2.7 Domain Decomposition Orderings

We next consider orderings based on domain decomposition. Figure 2.7a shows a standard 4-domain ordering, in which the points in each subdomain are ordered row-wise in such a way that within each domain the ordering proceeds from the outer corner inward. In Figure 2.7b, the subdomain points are ordered diagonally, so that vector or parallel properties of the diagonal ordering can be used within each subdomain. Duff and Meurant [1989] attribute these orderings within the subdomains to H. van der Vorst. In either case, the points in the separator set are numbered last and with these orderings the coefficient matrix takes the familiar arrowhead form

$$A = \begin{bmatrix} A_1 & & & & B_1^T \\ & A_2 & & & B_2^T \\ & & A_3 & & B_3^T \\ & & & A_4 & B_4^T \\ B_1 & B_2 & B_3 & B_4 & A_S \end{bmatrix} \tag{2.4}$$

Domain decomposition orderings have potentially good parallel properties, and were suggested by Farhat [1986] as a way to parallelize the SOR iteration. They can be used in the same way for SSOR or IC preconditioning. In particular, if there are p processors, it would be convenient to have p subdomains. Consider, for example, SSOR. In the solution of the corresponding lower triangular systems, each processor could solve one of the systems

$$L_i \mathbf{x}_i = \mathbf{d}_i \tag{2.5}$$

and form its contribution $B_i \mathbf{x}_i$ to the final system

$$L_S \mathbf{x}_S = \mathbf{b}_S - \sum_{i=1}^{p} B_i \mathbf{x}_i \tag{2.6}$$

Note that the additions in the right hand side of (2.6) require a fan-in, and the parallel solution of (2.6) is somewhat of a bottleneck.

In conjunction with SSOR or IC preconditioning, it is not necessary to have the separator set. If the separator points in Figure 2.7 are incorporated into the four subdomains and the subdomains are numbered counterclockwise, the coefficient matrix now takes the form

$$A = \begin{bmatrix} A_1 & B_1^T & 0 & B_4^T \\ B_1 & A_2 & B_2^T & 0 \\ 0 & B_2 & A_3 & B_3^T \\ B_4 & 0 & B_3 & A_4 \end{bmatrix} \tag{2.7}$$

where the A_i and B_i are not the same as in (2.4). For such non-separator domain decomposition

orderings, Duff and Meurant [1989] reported at least as good convergence results as for the natural ordering. (These orderings are called vdv1 and vdv2 in that paper, corresponding to row wise or diagonal ordering within the subdomains.) Unfortunately, without the separator sets, the natural parallelism of (2.4) is lost. However, it is still possible to work separately within each subdomain. For example, on the forward sweep of SSOR, SOR may be applied separately in each subdomain. When values from an adjacent subdomain are needed, the old ones are used. This is no longer SSOR on the whole domain, however; it is block Jacobi with respect to the subdomains and SSOR within each subdomain. Thus, the splitting of A for the forward SOR sweep is

$$A = \begin{bmatrix} D_1 - L_1 & & & \\ & D_2 - L_2 & & \\ & & D_3 - L_3 & \\ & & & D_4 - L_4 \end{bmatrix} - \begin{bmatrix} U_1 & -B_1^T & 0 & -B_4^T \\ -B_1 & U_2 & -B_2^T & 0 \\ 0 & -B_2 & U_3 & -B_3^T \\ -B_4 & 0 & -B_3 & U_4 \end{bmatrix} \tag{2.8}$$

where $D_i - L_i - U_i$ is the splitting of A_i into its diagonal and strictly triangular parts. Although this approach has perfect parallelism, the rate of convergence is likely to be inferior.

3. Condition Numbers

In this section, we collect some basic results on the condition number of the preconditioned matrix. If M is the preconditioning matrix, define the remainder matrix R by

$$A = M - R \tag{3.1}$$

Intuitively, the "smaller" R, the better M approximates A and the smaller the condition number of the preconditioned matrix, $M^{-1}A$, should be. More precisely,

$$M^{-1}A = I - M^{-1}R \tag{3.2}$$

and if the spectral radius $\rho(M^{-1}R)$ is small, then the eigenvalues of $M^{-1}A$ will all be close to 1, so that $M^{-1}A$ will be well-conditioned. Note that $M^{-1}R$ is the iteration matrix for a stationary linear iteration $\mathbf{x}^{k+1} = M^{-1}R\,\mathbf{x}^k + \mathbf{d}$ generated by the splitting (3.1), and rapid convergence of this iteration is equivalent to a better conditioned matrix $M^{-1}A$. We next give a more precise relation between the condition number of $M^{-1}A$ and the eigenvalues of $M^{-1}R$.

If A and M are symmetric positive definite, then the eigenvalues of $M^{-1}A$ are positive, and

$$\text{cond}(M^{-1}A) = \frac{\lambda_{\max}}{\lambda_{\min}} \tag{3.3}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the maximum and minimum eigenvalues of $M^{-1}A$. (Note that $M^{-1}A$ is not necessarily symmetric but it is similar to the preconditioned symmetric matrix $M^{-½}A\,M^{-½}$ and it is customary to write $\text{cond}(M^{-1}A)$ as the l_2 condition number of $M^{-½}A\,M^{-½}$). For the following theorem, recall that (3.1) is a regular splitting (Varga [1962]) if $M^{-1} \geq 0$ and $R \geq 0$, and a weak regular splitting (see, e.g. Ortega and Rheinboldt [1970]) if $M^{-1} \geq 0$ and $M^{-1}R \geq 0$.

Theorem 3.1. If A and M are symmetric positive definite and $\rho(M^{-1}R) < 1$, then

$$\text{cond}(M^{-1}A) = \frac{1 - \mu_{\min}}{1 - \mu_{\max}} \tag{3.4}$$

where $\mu_{\max}$ and $\mu_{\min}$ are the maximum and minimum eigenvalues of $M^{-1}R$. If, in addition, A is an M-matrix and (3.1) is a weak regular splitting, then

$$\text{cond}(M^{-1}A) = [1 + \rho(A^{-1}R)]\,(1 - \mu_{\min}) \tag{3.5}$$

Proof. By (3.2), the eigenvalues λ of $M^{-1}A$ and μ of $M^{-1}R$ are related by

$$\lambda = 1 - \mu$$

Since the eigenvalues of $M^{-1}A$ are all positive and since $\rho(M^{-1}R) < 1$, the eigenvalues of $M^{-1}R$

lie in the interval [0,1). Hence

$$\lambda_{min} = 1 - \mu_{max} \ , \ \lambda_{max} = 1 - \mu_{min}$$

and (3.4) follows. For the second part, by a theorem in Varga [1962] (proved for regular splittings but the same proof holds for weak regular splittings)

$$\rho(M^{-1}R) = \frac{\rho(A^{-1}R)}{1 + \rho(A^{-1}R)} \tag{3.6}$$

Thus,

$$1 - \mu_{max} = \frac{1}{1 + \rho(A^{-1}R)}$$

and (3.5) follows. This completes the proof.

One can make the estimate

$$1 - \mu_{min} \le 1 + |\mu_{min}| \le 1 + \rho(M^{-1}R) \tag{3.7}$$

and obtain the following corollary that was proved by Axelsson and Eijkhout [1989] under the weaker assumption that A and M are not necessarily symmetric but the eigenvalues of $M^{-1}R$ are real.

Corollary 3.1. Under the assumptions of Theorem 3.1,

$$\text{cond}(M^{-1}A) \le 1 + 2\rho(A^{-1}R) \tag{3.8}$$

The proof of this corollary follows from (3.7) and (3.5) by using (3.6) to obtain the identity

$$1 + \rho(M^{-1}R) = \frac{1 + 2\rho(A^{-1}R)}{1 + \rho(A^{-1}R)} \ .$$

Meijerink and van der Vorst [1977] showed that a no-fill ILU factorization of an M-matrix is a regular splitting. The same is true for the SSOR splitting with $\omega = 1$:

$$M = (D - L)\, D^{-1} (D - L^T), \quad R = L D^{-1} L^T \tag{3.9}$$

where L is the strictly lower triangular part of A and D the diagonal part. Thus, if A is symmetric positive definite, Theorem 3.1 and Corollary 3.1 apply to both of these preconditioners. Moreover, for (3.9) we can obtain a much sharper result.

Corollary 3.2. If the conditions of Theorem 3.1 hold and R is positive semidefinite, then

$$\text{cond}(M^{-1}A) \le 1 + \rho(A^{-1}R) \tag{3.10}$$

and if R is singular

$$\text{cond}(M^{-1}A) = 1 + \rho(A^{-1}R) \tag{3.11}$$

In particular, (3.11) holds for (3.9)

Proof: The eigenvalues of R, and hence $M^{-1}R$, are nonnegative so that $1 - \mu_{min} \le 1$; thus, (3.10) follows from (3.5). If R is singular, then $\mu_{min} = 0$ so that (3.11) is (3.5) in this case. Finally, since L in (3.9) is strictly lower triangular, R is singular.

We note that Corollary 3.2 does not apply to incomplete Cholesky factorization, even when $R \ge 0$, since R need not be positive semi-definite. Moreover, the weaker but sufficient condition that the eigenvalues of $M^{-1}R$ be non-negative does not necessarily hold either, as is seen by the simplest example of the Poisson equation and no-fill incomplete factorization.

4. Effect of Ordering

We now wish to consider the effects of ordering on the rate of convergence. Let A_N be the coefficient matrix in the natural ordering and $A_C = PA_NP^T$, where P is some permutation matrix. If $A_N = M_N - R_N$ and $A_C = M_C - R_C$ are the corresponding splittings (3.1), then

$$A_C^{-1}R_C = (PA_NP^T)^{-1}\, R_C = PA_N^{-1}P^T R_C P\, P^T \tag{4.1}$$

Thus, $A_C^{-1}R_C$ and and $A_N^{-1}P^TR_CP$ are permutationally similar so that

$$\rho(A_C^{-1}R_C) = \rho(A_N^{-1}P^TR_CP), \quad \|A_C^{-1}R_C\|_2 = \|A_N^{-1}P^TR_CP\|_2 \tag{4.2}$$

Therefore, the difference in the estimates (3.8)

$$\text{cond}\,(M_N^{-1}A_N) \le 1 + 2\rho(A_N^{-1}R_N) \le 1 + 2\|A_N^{-1}\|_2\,\|R_N\|_2 \tag{4.3}$$

$$\text{cond}\,(M_C^{-1}A_C) \le 1 + 2\rho(A_N^{-1}P^TR_CP) \le 1 + 2\|A_N^{-1}\|_2\,\|R_C\|_2 \tag{4.4}$$

depends only on the difference between $\|R_N\|_2$ and $\|R_C\|_2$ in the case of the norm estimates. However, these norm estimates may be overly pessimistic and the more critical factors are $\rho(A_N^{-1}R_N)$ and $\rho(A_N^{-1}P^TR_CP)$. This is especially the case when Corollary 3.2 applies so that we have the exact condition numbers.

We first consider under what conditions

$$\rho(A_N^{-1}R_N) = \rho(A_N^{-1}P^TR_CP) \tag{4.5}$$

Clearly, a sufficient condition is

$$R_N = P^TR_CP \tag{4.6}$$

which is equivalent to

$$M_N = P^TM_CP \tag{4.7}$$

and we will use (4.7) and (4.6) interchangeably. We note that, in fact, the eigenvalues of $M_N^{-1}A_N$ and $M_C^{-1}A_C$ are the same if (4.7) holds for any non-singular matrix P since

$$M_N^{-1}A_N = (P^TM_CP)^{-1}\,(P^TA_CP) = P^{-1}M_C^{-1}A_CP$$

so that $M_N^{-1}A_N$ and $M_C^{-1}A_C$ are similar.

We next consider an example of when (4.6) holds. For the SSOR splitting (3.9), where $A_N = D_N - L_N - L_N^T$ and $A_C = D_C - L_C - L_C^T$, (4.6) may be written as

$$L_N D_N^{-1} L_N^T = P^T L_C P \, P^T D_C^{-1} P \, P^T L_C^T P \tag{4.8}$$

Since $D_N = P^T D_C P$, (4.8) holds if

$$L_N = P^T L_C P \tag{4.9}$$

For incomplete Cholesky factorization, we have a similar condition; in this case, if $\tilde{L}_N$ and $\tilde{L}_C$ are the factors, then

$$R_N = \tilde{L}_N \tilde{L}_N^T - A_N, \quad R_C = \tilde{L}_C \tilde{L}_C^T - A_C$$

so that (4.6) is

$$\tilde{L}_N \tilde{L}_N^T - A_N = P^T \tilde{L}_C P \, P^T \tilde{L}_C^T P - P^T A_C P$$

Thus a sufficient condition for (4.6) to hold is

$$\tilde{L}_N = P^T \tilde{L}_C P \tag{4.10}$$

For a no-fill incomplete factorization, $\tilde{L}_N$ will have the same nonzero off-diagonal structure as L_N, and similarly for $\tilde{L}_C$ and L_C. Thus, (4.10) will hold if (4.9) does. Young [1971] has a defined a permutation matrix P to be *non-migratory* if (4.9) holds, and has shown that the permutation matrix that transforms the natural ordering to the diagonal ordering of Figure 1.1 is non-migratory. Hence, these ordering are equivalent (as is intuitively clear).

We can characterize the relation (4.9) in the following way, in terms of the Gauss-Seidel iterations

$$D_N \mathbf{u}_N^{k+1} - L_N \mathbf{u}_N^{k+1} - U_N \mathbf{u}_N^k = \mathbf{b}_N \tag{4.11a}$$

$$D_C \mathbf{u}_C^{k+1} - L_C \, \mathbf{u}_C^{k+1} - U_C \mathbf{u}_C^k = \mathbf{b}_C \tag{4.11b}$$

We can rewrite (4.11b) as

$$P^T D_C P \, P^T \mathbf{u}_C^{k+1} - P^T L_C P \, P^T \mathbf{u}_C^{k+1} - P^T U_C P \, P^T \mathbf{u}_c^k = P^T \mathbf{b}_C \tag{4.12}$$

Thus, if (4.9) holds (and, consequently, $U_N = P^T U_C P$ also), then (4.12) becomes

$$D_N P^T \mathbf{u}_C^{k+1} - L_N P^T \mathbf{u}_C^{k+1} - U_N P^T \mathbf{u}_C^k = \mathbf{b}_N \tag{4.13}$$

Therefore, provided that $\mathbf{u}_N^0 = P^T \mathbf{u}_C^0$, (4.11a) and (4.13) generate exactly the same sequences $\mathbf{u}_N^k$ and $P^T \mathbf{u}_C^k$ so that

$$\mathbf{u}_C^k = P\, \mathbf{u}_N^k \tag{4.14}$$

On the other hand, if (4.9) does not hold, (4.11a) and (4.11b) will not generate sequences for which (4.14) holds, and the rates of convergence may be much different. Duff and Meurant [1989] tabulated the maximum element of R and the Frobenius norm of R (= $(\sum r_{ij}^2)^{1/2}$) for several different orderings on some two-dimensional Poisson-type model problems of the form (1.2) using ICCG. While these numbers are interesting, they are only qualitatively related to differences in the rates of convergence. The key quantity is $\rho(A^{-1}R)$, as discussed in the previous section. Figure 4.1 compares these quantities for SSOR ($\omega = 1$) preconditioning on the problem (1.2) for 7×7 and 15×15 grids. In this case, (3.11) holds and thus $1 + \rho(A^{-1}R)$ gives the exact condition number.

	$1 + \rho(A^{-1}R)$	$\max\|r_{ij}\|$	$\|R\|_F$
Nat(7×7)	3.99	.125	.89
RB(7×7)	6.83	.25	2.56
Nat(15×15)	13.7	.125	4.7
RB(15×15)	23.6	.25	13.7

Figure 4.1. SSOR preconditioning

As seen in Figure 4.1, $\max |r_{ij}|$ gives very little indication of the difference in the condition numbers. $\|R\|_F$ gives a better indication in that it increases as the condition number increases, but it is of marginal value in ascertaining the size of the condition number.

Even though R by itself does not determine the condition number accurately, it is certainly a key factor and we would like to understand how and why R changes for different orderings. In the sequel, we will assume that A arises from the 5-point discretization of the Poisson-type problem (1.2) on a $N \times N$ grid, and that M has the form

$$M = (\overline{D} - L)\,\overline{D}^{-1}(\overline{D} - L^T) \tag{4.15}$$

where $\overline{D}$ is diagonal and L is the strictly lower triangular part of A. This will be the case for no-fill Cholesky factorization (van der Vorst [1982]) or symmetric Gauss-Seidel (SGS), but not for SSOR with $\omega \neq 1$. Then

$$R = M - A = \overline{D} - D + L\overline{D}^{-1}L^T \tag{4.16}$$

where D is the main diagonal of A. Note that in the case of SGS, $\overline{D} = D$ so that $R = LD^{-1}L^T$.

We next obtain the structure of R by performing the multiplications $L\overline{D}^{-1}L^T$ in the following way. The ith row of $L\overline{D}^{-1}L^T$ is $e_i^T L\overline{D}^{-1}L^T$, where e_i is the ith unit vector. In the natural ordering, the ith row of L has two non-zero elements and may be written as

$$e_i^T L = a_{i,i-1}\, e_{i-1}^T + a_{i,i-N}\, e_{i-N}^T \tag{4.17}$$

where the a_{ij} are off-diagonal elements of A. Then

$$e_i^T L\overline{D}^{-1}L^T = a_{i,i-1}\,\overline{d}_{i-1}^{-1} e_{i-1}^T L^T + a_{i,i-N}\,\overline{d}_{i-N}^{-1} e_{i-N}^T L^T$$

Now

$$e_{i-1}^T L^T = a_{i,i-1}\, e_i^T + a_{i+N-1,i-1}\, e_{i+N-1}^T$$

and

$$e_{i-N}^T L^T = a_{i-N+1,i-N}\, e_{i-N+1}^T + a_{i,i-N}\, e_i^T$$

Thus

$$e_i^T L\overline{D}^{-1}L^T = a_{i,i-1}^2 \overline{d}_{i-1}^{-1} e_i^T + a_{i,i-1}a_{i+N-1,i-1}\overline{d}_{i-1}^{-1} e_{i+N+1}^T \tag{4.18}$$

$$+ a_{i,i-N}a_{i-N+1,i-N}\overline{d}_{i-N}^{-1} e_{i-N+1}^T + a_{i,i-N}^2 \overline{d}_{i-N}^{-1} e_i^T$$

The expression (4.18) for the ith row of $L\overline{D}^{-1}L^T$ shows that R is a three-diagonal matrix, in which the two off-diagonals are at a distance N−1 from the main diagonal; that is, they are shifted one position in, relative to the outer-most diagonals of A itself. The magnitudes of the elements of R are given by (4.18) and (4.16):

$$\text{off–diagonals: } a_{i,i-1}a_{i+N-1,i-1}\overline{d}_{i-1}^{-1},\ a_{i,i-N}a_{i-N+1}\overline{d}_{i-N}^{-1} \tag{4.19a}$$

$$\text{diagonal: } a_{i,i-1}^2\overline{d}_{i-1}^{-1} + a_{i,i-N}^2\overline{d}_{i-N}^{-1} + \overline{d}_i - a_{ii} \tag{4.19b}$$

Consider next the red/black ordering in which

$$A = \begin{bmatrix} D_1 & C^T \\ C & D_2 \end{bmatrix},\quad L = \begin{bmatrix} 0 & 0 \\ C & 0 \end{bmatrix}$$

so that

$$R = \overline{D} - D + \begin{bmatrix} 0 & 0 \\ 0 & C\overline{D}_2^{-1}C^T \end{bmatrix}$$

Elman and Golub [1988] have computed the reduced system $C\,D_2^{-1}C^T$ for a somewhat different equation but their general conclusions hold here also. In particular, $C\,\overline{D}_2^{-1}C^T$ is a nine-diagonal matrix corresponding to the stencil shown in Figure 4.2. (Note that the points of the original five-point stencil, except for the center point, do not appear in the stencil for R. This would not be the case, however, for SSOR with $\omega \neq 1$.)

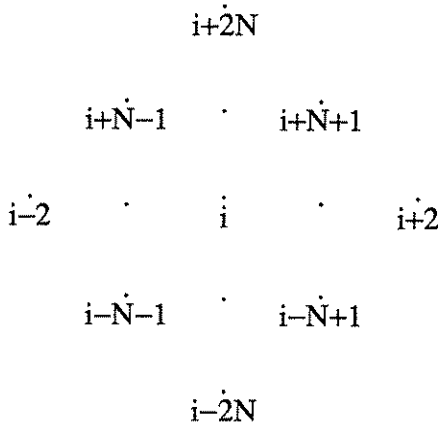


Figure 4.2 Nine-point Stencil for R: Red/black ordering

The number of non-zero elements in $C\overline{D}^{-1}C^T$ is $0(\frac{9}{2}N^2 - 8N)$. Thus, R has either this number of non-zero elements or $0(\frac{N^2}{2})$ more in case $D \neq D_1$. In contrast, R for the natural ordering has $0(3N^2 - 2N)$ non-zero elements, so that the red/black remainder matrix has about 50 percent more non-zero elements, for large N. Moreover, the magnitudes of the non-zero elements in the red/black remainder matrix are larger, as we now discuss.

The elements of R for the natural ordering are given by (4.19). For the red/black ordering, we will not give the elements precisely but will indicate their magnitude as follows. The elements of $\overline{D}^{-1}L^T$ are of the form $d^{-1}a$, where a is an element of A and d is an element of $\overline{D}$. When $\overline{D}^{-1}L^T$ is multiplied by $e_i^T L$, the elements of $\overline{D}^{-1}L^T$ are multiplied by another element of A to give terms of the form $d^{-1}aa$. The important thing is that some of these terms may add together. In particular, the diagonal elements of $L\overline{D}^{-1}L^T$ will consist of a sum of four terms. Elements in the ith row corresponding to the four grid points $i\pm N\pm 1$ in Figure 4.2 will consist of a sum of two terms, and the elements corresponding to the grid points $i\pm 2N$ and $i\pm 2$ consist of a single term. Comparing this with (4.19), we would expect that, roughly, the diagonal elements of R corresponding to $L\overline{D}^{-1}L^T$ for the red/black ordering will be twice as large as the diagonal

elements of R for the natural ordering, and half of the off-diagonal elements will also be twice as large. If we assume that the elements of the R for the natural ordering all have the same magnitude r, then the Frobenius norm is $\|R\|_{F,Nat} \doteq \sqrt{3}\ Nr$, based on the estimate of $0(3N^2)$ elements in R. For the red/black ordering, we would have $0(\frac{1}{2}N^2)$ elements with magnitude 4r, $0(2N^2)$ with magnitude 2r and $0(2N^2)$ with magnitude r. Thus, $\|R\|_{F,RB} \doteq \sqrt{18}\ Nr$ so that the ratio of the Frobenius norms is about 2.5. This very rough estimate agrees quite well with the data in Figure 4.1 for SSOR preconditioning, in which the ratios are about 3 and with three of the four problems in Duff and Meurant [1989] for ICCG for which the ratios are about 2. (The fourth problem has very strong anisotropy and the ratio is 43).

We comment on the difference of SOR and SSOR in terms of the remainder matrix. Consider the Gauss-Seidel splitting

$$A = (D - L) - L^T$$

Here, in analogy to the splitting $A = M - R$, $D - L$ corresponds to M and L^T corresponds to R. In terms of the natural and red/black orderings, L_N^T and L_{RB}^T have the same number of elements and, in the case of Poisson's equation, exactly the same elements, just in different positions. Hence, the ordering has no effect on the "size" of R. For SSOR ($\omega = 1$), however, the remainder matrices

$$(D_N - L_N)\, D_N^{-1} (D_N - L_N^T), \quad (D_{RB} - L_{RB})\, D_{RB}^{-1} (D_{RB} - L_{RB}^T)$$

are very different, as we have just seen, with the red/black remainder matrix having many more, as well as larger, elements.

We consider one more class of orderings, those based on domain decomposition as discussed in Section 2. Consider first the ordering of Figure 2.7a with four domains and a

separator set, so that the coefficient matrix A is given by (2.4). Then the strictly lower triangular part of A is

$$L = \begin{bmatrix} L_1 & & & & \\ & L_2 & & & \\ & & L_3 & & \\ & & & L_4 & \\ B_1 & B_2 & B_3 & B_4 & L_5 \end{bmatrix}$$

when L_i is the strictly lower triangular part of A_i. The diagonal blocks of $L\,\overline{D}^{-1}L^T$ are then

$$L_i\overline{D}_i^{-1}L_i^T,\ i = 1, \cdots, 4,\quad L_5\overline{D}_5^{-1}L_5^T + \sum_{i=1}^{4} B_i\overline{D}_i^{-1}B_i^T \tag{4.20}$$

and the off-diagonal blocks are

$$L_i\overline{D}_i^{-1}B_i^T,\ i = 1, \cdots, 4,\quad B_i\overline{D}_i^{-1}L_i^T,\ i = 1, \cdots, 4 \tag{4.21}$$

For simplicity, we will assume that each subdomain has $N_1 \times N_1$ grid points so that $N = 2N_1 + 1$, and there are $4N_1 + 1$ points in the separator set.

The matrices $L_i\overline{D}_i^{-1}L_i^T$, $i = 1, \cdots, 4$, have the three-diagonal structure for the natural order, as discussed earlier. The matrix L_5 has the form

$$L_5 = \begin{bmatrix} L_{5,1} & & & & \\ & L_{5,2} & & & \\ & & L_{5,3} & & \\ & & & L_{5,4} & \\ v_1^T & v_2^T & v_3^T & v_4^T & 0 \end{bmatrix}$$

where each $L_{5,i}$ is lower bidiagonal (with zero main diagonal) and the v_i are column vectors with a non-zero element in the last position. It follows that $L_5\overline{D}_5^{-1}L_5^T$ is diagonal: each diagonal element is zero or of the form $d^{-1}aa$ except the last element, which is a sum of four such terms.

As before, a denotes some element of A, and d some element of $\overline{D}$.

The matrix B_i^T gives the connections of the points of domain i with the separator points and the only non-zero rows correspond to interior grid points adjacent to the separator points. Consider B_1; the situation for the other B_i is similar. The non-zero rows of B_1 corresponding to points along the vertical and horizontal separator sets are

$$a\,\mathbf{e}_i^T,\ i = N_1, 2N_1, \cdots, N_1^2,\quad a\,\mathbf{e}_i^T,\ i = N_1^2, N_1^2 - 1, \cdots, N_1^2 - N_1 + 1 \tag{4.22}$$

where the element a of A is different in each case. Thus, B_1 may be expressed as a sum of rank one matrices of the form $a\,\mathbf{e}_j\,\mathbf{e}_i^T$, where i is one of the indices of (4.22) and j denotes the position of the corresponding row in B_1. Therefore, $B_1\,\overline{D}_1^{-1}B_1^T$ is a sum of terms of the form

$$a\,a\,d^{-1}\mathbf{e}_{j_1}\,\mathbf{e}_{i_1}^T\mathbf{e}_{i_2}\,\mathbf{e}_{j_2}^T$$

and these matrices are non-zero only when $i_1 = i_2 = i$, where i is one of the indices of (4.22). Hence, $B_1\overline{D}_1^{-1}B_1^T$ consists of $2N_1$ non-zero elements of the form $a\,a\,d^{-1}$. Likewise, the other matrices $B_i\overline{D}_i^{-1}B_i^T$ have $2N_1$ non-zero elements.

Finally, consider the matrix $L_1\overline{D}_1^{-1}B_1^T$. The ith row of $L_1\overline{D}_1^{-1}$ is, in general, of the form

$$a\,d^{-1}\mathbf{e}_{i-1}^T + a\,d^{-1}\mathbf{e}_{i-N}^T,$$

Thus, the ith row of $L_1\overline{D}_1^{-1}B_1^T$ is a linear combination of the (i–1)st and (i–N_1)th rows of B_1^T. But both of these rows cannot be simultaneously non-zero so that $L_1\overline{D}_1^{-1}B_1^T$ has no more non-zero elements that B_1 itself. Similarly, the other $L_i\overline{D}_i^{-1}B_i^T$ have no more non-zero elements than B_i.

If we add the number of non-zero elements of each submatrix in $L\overline{D}^{-1}L^T$, we have at most

$$4 \cdot 3N_1^2 + (4N_1 + 4 \cdot 2N_1) + 8 \cdot 2N_1 = 12\,N_1^2 + 28N_1 \tag{4.23}$$

non-zero elements. The first term in (4.23) gives the number of non-zero elements in the first four matrices of (4.20), the second term is for the last matrix of (4.20), and the third term is for the eight matrices of (4.21). We have previously shown that the remainder matrix for the natural ordering has approximately $3N^2$ non-zero elements. Since $N = 2N_1 + 1$, this gives $12N_1^2 + 12N_1 + 1$ which is only slightly smaller than (4.23) for large N. Thus, the remainder matrices for the natural and domain decomposition orderings have roughly the same number of non-zero elements. Although the above discussion has only been for the case of four subdomains, one would expect that the same conclusions would hold in more generality.

Suppose, next, that we remove the separator sets so that A has the form (2.7). In this case

$$L = \begin{bmatrix} L_1 & & & \\ B_1 & L_2 & & \\ 0 & B_2 & L_3 & \\ B_4 & 0 & B_3 & L_4 \end{bmatrix}$$

and the remainder matrix has the following submatrices

$$L_1\overline{D}_1^{-1}L_1^T,\ L_i\overline{D}_i^{-1}L_i^T + B_{i-1}\overline{D}_{i-1}^{-1}B_{i-1}^T,\ i = 2, 3,\ L_4\overline{D}_4^{-1}L_4^T + B_4\overline{D}_4^{-1}B_4^T + B_3\overline{D}_3^{-1}B_3^T \tag{4.24}$$

$$L_i\overline{D}_i^{-1}B_i^T,\ i = 1, 2, 3,\ L_1\overline{D}_1^{-1}B_4^T,\ B_1\overline{D}_1^{-1}B_4^T \tag{4.25}$$

plus the transposes of the matrices in (4.25). The $L_i\overline{D}_i^{-1}L_i^T$ have the same natural order structure as before, and the B_i have the same general structure as in the separator case since they provide the connections along the horizontal and vertical interfaces of the subdomains. A rough estimate of the non-zero elements in $L\overline{D}^{-1}L^T$ is then

$$3N_1^2 + 2(3N_1^2 + 2N_1) + 3N_1^2 + 2 \cdot 2N_1 + 10 \cdot 2N_1 = 12N_1^2 + 28N_1 \tag{4.26}$$

Since now $N = 2N_1$, $3N^2 = 12N_1^2$ is an estimate of the number of non-zero elements in the natural order remainder matrix. Again, for large N, this is very close to the estimate (4.26). This

tends to explain why Duff and Meurant [1989] obtained essentially the same number of iterations for their domain decomposition orderings (the "vdv" orderings) as for the natural ordering.

References

Adams, L. and Jordan, H. [1985]. "Is SOR Color-Blind?" *SIAM J. Sci. Stat. Comput.* 7, pp. 490-506.

Ashcraft, C. and Grimes, R. [1988]. "On Vectorizing Incomplete Factorization and SSOR Preconditioners." *Siam J. Sci. Stat. Comput.* 9, pp. 122-151.

Axelsson, O. and Eijkhout, V. [1989]." Vectorizable Preconditioners for Elliptic Difference Equations in Three Space Dimensions." *J. Comp. Appl. Math.* 27, pp. 299-321.

Chan, T., Kuo, C.C., and Tong, C. [1989]. Parallel Elliptic Preconditioners: Fourier Analysis and Performance on the Connection Machine. *Comput. Phys. Comm.* 53, pp. 237-252.

Cuthill, E. and McKee, J. [1969]. "Reducing the Bandwidth ofSparse Symmetric Matrices." *Proc. 24th Nat. Conf. ACM,* pp. 157-172.

Duff, E. and Meurant, G. [1989]. "The Effect of Ordering on Preconditioned Conjugate Gradients." *BIT* 29, pp. 635-657.

Eisenstat, S. [1981]. "Efficient Implementation of a Class of Conjugate Gradient Methods." *SIAM J. Sci. Stat. Comput.* 2, pp. 1-4.

Elman, H. and Agron, E [1989]. "Ordering Techniques for the Preconditioned Conjugate Gradient-Method on Parallel Computers." *Comput. Phys. Comm.* 53, pp. 253-269.

Elman, H. and Golub, G. [1988]. "Iterative Methods for Cyclically Reduced Non-Self-Adjoint Linear Systems." *Computer Science Rep. CS-TR-2145.* Univ. Maryland.

Ericksen, J. [1972]. "Iterative and Direct Methods for Solving Poisson's Equation and their Applicability to ILLIAC IV." *Tech Rep. 60,* Center for Advanced Computation, Univ. ofIllinois at Urbana-Champaign.

Farhat, C. [1986]. "Multiprocessors in Computational Mechanics." *Ph.D. Thesis,* Civil Engineering, Univ. California-Berkeley.

Greenbaum, A [1986]. "Solving Sparse Triangular Linear Systems using Fortran with Parallel Extensions on the NYU Ultracomputer Prototype." *Ultracomputer Note* 99, New York

University, New York.

Hageman, L. and Young, D. [1981]. "Applied Iterative Methods." Academic Press, New York.

Harrar, D. and Ortega, J. [1990]. "Solution of Three-Dimensional Generalized Poisson Equations on Vector Computers." *Iterative Methods for Large Linear Systems*, D. Kincaid and L. Hayes, Eds., Academic Press, New York, pp. 173-191.

Hayes, L. [1978]. "Timing Analysis of Standard Iterative Methods on a Pipeline Computer." *Rep. CNA-136*, Center Numer. Analysis., U. Texas at Austin.

Lambiotte, J. [1975]. "The Solution of Linear Systems of Equations on a Vector Computer." *Ph.D. Thesis*, Applied Mathematics, U. Virginia.

Madsen, N., Rodrigue, G. and Karush, J. [1976]. "Matrix Multiplication by Diagonals on a Vector/Parallel Processor." *Inf. Proc. Letters 5*, pp. 41-45.

Meijerink, J. and van der Vorst, H. [1977]. "An Iterative Solution for Linear Systems of which the Coefficient Matrix is a Symmetric M-Matrix." *Math. Comput.* 31, pp. 148-162.

Melhem, R. [1986]. "Toward Efficient Implementation of Preconditioned Conjugate Gradient Methods on Vector Supercomputers." *Int. J. Supercomputer Appl.* 2, pp. 70-98.

Ortega, J. [1988]. "Introduction to Parallel and Vector Solution of Linear Systems." Plenum Pub. Co., New York.

Ortega, J. and Rheinboldt, W. [1970]. "Iterative Solution ofNonlinear Equations in Several Variables." Academic Press, New York.

Poole, E. and Ortega, J. [1987]. "Multicolor ICCG Methods for Vector Computers." *SIAM J. Numer. Anal.* 24, pp. 1394-1418.

Saltz, J. [1990]. "Aggregation Methods for Solving Sparse Triangular Systems on Multiprocessors." *SIAM J. Sci. Stat. Comput.* 11, pp. 123-145.

Schlichting, J. and van der Vorst, H. [1989]. "Solving 3D Block Bidiagonal Linear Systems on Vector Computers." *J. Comp. Appl. Math.* 27, pp. 323-330.

Stotland, S. [1990]. "Vectorizing the Conjugate Gradient Method with Preconditioning using the Diagonal Ordering." *Master's Project*, Applied Mathematics, U. Virginia.

van der Vorst, H. [1982]. "A Vectorizable Variant of Some ICCG Methods." *SIAM J. Sci. Stat. Comput.* 3, pp. 350-356.

van der Vorst, H. [1983]. "On the Vectorization of Some Simple ICCG Methods." First Int. Conf. Vector and Parallel Computation in Scientific Applications, Paris.

van der Vorst, H. [1987]. "Large Tridiagonal and Block Tridiagonal Linear Systems on Vector and Parallel Computers."*Parallel Computing* 5, pp. 45-54.

van der Vorst, H. [1989a]. "High Performance Preconditioning." *SIAM J. Sci. Stat. Comput.* 10, pp. 1174-1185.

van der Vorst, H. [1989b]. "ICCG and Related Methods for 3D Problems on Vector Computers." *Comp. Phys. Comm.* 53, pp. 223-235.

Varga, R. [1962]. "Matrix Iterative Analysis." Prentice-Hall, Englewood Cliffs, New Jersey.

Young, D. [1971]. "Iterative Solution of Large Linear Systems." Academic Press, New York.