```
      subroutine prec3s( x, px, a2, a3, a4, nx, ny, nz, rkr )
      implicit double precision (a-h,o-z)
      dimension x(nx,ny,nz), px(nx,ny,nz), a2(nx,ny,nz),
      a3(nx,ny,nz),
     &          a4(nx,ny,nz)
      px(1,1,1) = x(1,1,1)
      do i = 2, nx
         px(i,1,1) = x(i,1,1) - a2(i-1,1,1) * px(i-1,1,1)
      enddo
      do j = 2, ny
         do i = 1, nx
            px(i,j,1) = x(i,j,1) - a2(i-1,j,1) * px(i-1,j,1)
     &                   - a3(i,j-1,1) * px(i,j-1,1)
         enddo
      enddo
      do k = 2, nz
         do j = 1, ny
            do i = 1, nx
               px(i,j,k) = x(i,j,k) - a2(i-1,j,k) * px(i-1,j,k)
     &                     - a3(i,j-1,k) * px(i,j-1,k)
     &                     - a4(i,j,k-1) * px(i,j,k-1)
            enddo
         enddo
      enddo
*
*     Lower triangular part has been solved
*
*     Computation of (r,w), where w is the preconditioned residual
*
      rkr = ddot( nx * ny * nz, px, 1, px, 1 )
      do i = nx-1, 1, -1
           px(i,ny,nz) = px(i,ny,nz) - a2(i,ny,nz) * px(i+1,ny,nz)
      enddo
      do j = ny-1, 1, -1
         do i = nx, 1, -1
            px(i,j,nz) = px(i,j,nz) - a2(i,j,nz) * px(i+1,j,nz)
     &                   - a3(i,j,nz) * px(i,j+1,nz)
         enddo
      enddo
      do k = nz-1, 1, -1
         do j = ny, 1, -1
            do i = nx, 1, -1
               px(i,j,k) = px(i,j,k) - a2(i,j,k) * px(i+1,j,k)
     &                     - a3(i,j,k) * px(i,j+1,k)
     &                     - a4(i,j,k) * px(i,j,k+1)
            enddo
         enddo
      enddo
*
*     Diagonal and upper triangular part has been solved
*
      return
      end
```

```
      subroutine precd( px, nxny, a2, a3, nx, left )
      implicit double precision (a-h,o-z)
      dimension px(*), a2(*), a3(*)
      logical left
*
*     Diagonalwise computation in preconditioning over x,y-plane
*     ([2]: Chapter 7.3.4)
*
      ny = nxny / nx
      if (.not.left) go to 15
      do 10 ii = 2, nx + ny - 1
         jjmin = max( 1, ii - nx + 1 )
         jjmax = min( ii, ny )
         is = (jjmin - 1) * (nx - 1) + ii
c$dir no_recurrence
         do 5 j = jjmin, jjmax
            px(is) = px(is) - a2(is-1) * px(is-1) -
     &               a3(is-nx) * px(is-nx)
            is = is + nx - 1
    5    continue
   10 continue
      return
*
   15 continue
      do 30 ii = nx + ny - 1, 2, -1
         jjmin = max( 1, ii - nx + 1 )
         jjmax = min( ii, ny )
         is = (jjmin - 1) * (nx - 1) + ii
c$dir no_recurrence
         do 20 j = jjmin, jjmax
            px(is) = px(is) - a2(is) * px(is+1) -
     &               a3(is) * px(is+nx)
            is = is + nx - 1
   20    continue
   30 continue
      px(1) = px(1) - a2(1) * px(2) - a3(1) * px(nx+1)
      return
      end
```

C. Partly Vectorized Standard ICCG

The subroutine `prec3s` can be partly vectorized. This approach leads to code as in `prec3p` for which the compiler may detect that there is a recurrence left only in the first index (i.e., in the x-direction of the grid). The call to `prec3s` in the calling routine `iccg` should be changed accordingly in a call to `prec3p`.

The motivation for providing this version is the following. In some situations, when the matrix A has a different nonzero structure, vectorization by different ordering (as in `prec3d`) is not possible. In such cases partial vectorization might be appropriate. Comparing the performances of `prec3p` and `prec3s` then will give an idea of the effects of partial vectorization that one may expect in more general situations.

B. Diagonal-Order ICCG

For the 7-point stencil, the incomplete Cholesky preconditioner can be vectorized by reordering the unknowns diagonalwise over x, y planes of the grid. The subroutine `prec3d` carries out the preconditioning step over the whole grid and leaves the parts corresponding to x, y planes to the subroutine `precd`. These codes contain comment directives (in our example for Convex compilers) that instruct the compiler to ignore the apparent recurrence. These directives should be adjusted to the respective computer.

```
      subroutine prec3d( x, px, n, nx, ny, nz, rkr )
      implicit double precision (a-h,o-z)
      dimension x(n), px(n)
      common /matrix/ a(1061208,10)
      logical left
*
*     Preconditioning for ICCG
*     Vectorization by diagonalwise ordering of computations
*             in (x,y) plane ([2]: Chapter 7.3.4)
*
      nxny = nx * ny
*
*     Solving the lower triangular part of the preconditioner
*
      left= .true.
      do 10 i = 1, nxny
         px(i) = x(i)
   10 continue
      call precd( px(1), nxny, a(1,2), a(1,3), nx, left )
      do 30 j = 2, nz
         jl = (j - 1) * nxny + 1
         ju= j * nxny
c$dir no_recurrence
         do 20 i = jl, ju
            px(i) = x(i) - a(i-nxny,4) * px(i-nxny)
   20    continue
         call precd( px(jl), nxny, a(jl,2), a(jl,3), nx, left )
   30 continue
      rkr = ddot( n, px, 1, px, 1 )
      left = .false.
*
*     Solving the upper triangular part of the preconditioning
*
      jl = (nz - 1) * nxny + 1
      call precd( px(jl), nxny, a(jl,2), a(jl,3), nx, left )
      do 50 j = nz-1, 1, -1
         jl = (j - 1) * nxny + 1
         ju = j * nxny
c$dir no_recurrence
         do 40 i = jl, ju
            px(i) = px(i) - a(i,4) * px(i+nxny)
   40    continue
         call precd( px(jl), nxny, a(jl,2), a(jl,3), nx, left )
   50 continue
      return
      end
```

Appendix: Details of the Relevant Parts of the Software

A. Standard ICCG

The standard incomplete Cholesky preconditioning is carried out by subroutine `prec3s`.

```
      subroutine prec3s( x, px, n, nxny, nx, rkr )
      implicit double precision (a-h,o-z)
      dimension x(n), px(n)
      common /matrix/ a(1061208,10)
*
*     Standard preconditioning ICCG(0)
*
      px(1) = x(1)
      do 10 i = 2, nx
         px(i) = x(i) - a(i-1,2) * px(i-1)
   10 continue
      nx1 = nx + 1
      do 20 i= nx1, nxny
         px(i) = x(i) - a(i-1,2) * px(i-1) - a(i-nx,3) * px(i-nx)
   20 continue
      nxny1 = nxny + 1
      do 25 i = nxny1, n
         px(i) = x(i) - a(i-1,2) * px(i-1) - a(i-nx,3) * px(i-nx) -
     &         a(i-nxny,4) * px(i-nxny)
   25 continue
*
*     Lower triangular part has been solved
*
*     Computation of (r,w), where w is the preconditioned residual
*
      rkr = ddot( n, px, 1, px, 1 )
      ij = n - 1
      do 30 i = 2, nx
         px(ij) = px(ij) - a(ij,2) * px(ij+1)
         ij = ij - 1
   30 continue
      do 40 i = nx1, nxny
         px(ij) = px(ij) - (a(ij,2) * px(ij+1) + a(ij,3) * px(ij+nx))
         ij = ij - 1
   40 continue
      do 50 i = nxny1, n
         px(ij) = px(ij) - (a(ij,2) * px(ij+1) + a(ij,3) * px(ij+nx)
         +
     &            a(ij,4) * px(ij+nxny))
         ij=ij-1
   50 continue
*
*     Diagonal and upper triangular part has been solved
*
      return
      end
```

[4] R. W. Freund, G. H. Golub, and N. M. Nachtigal. *Iterative solution of linear systems.* Numerical Analysis Report Stanford University, NA-91-05, 1991.

[5] H. A. van der Vorst. High-performance preconditioning. *SIAM J. Sci. Statist. Comput.*, 10:1174–1185, 1989.

4.2 Rules for Running the Tests

The software intentionally has been kept simple so that it will be easy for an experienced programmer to adapt the program, or parts of it, to a specific architecture with only a modest effort. In running the tests, the user is allowed to reorganize the CG algorithm or parts of it (e.g., grouping updates or combining inner-products with other operations); calls to the BLAS may be replaced by optimized code, and compiler optimization directives may be added. However, the user may not change the basic algorithm.

5 Conclusions

The state of optimization for sparse matrix algorithms is much different from that of dense problems. This fact can be readily verified from the results in Table 1. For dense problems one can typically exploit features of the architecture by using block versions of the algorithm (for cache and distributed memory), by combining successive updates (vector register machines), or by exploiting both methods. In fact, many dense linear algebra algorithms can be organized so that $\mathcal{O}(n)$ arithmetic operations can be done per data transfer from main memory to the fastest memory. In contrast, the ratio of computational work to data transfer is usually rather low ($\mathcal{O}(1)$) for iterative sparse matrix solvers, and the vectors are usually much too long to be kept in any local memory. Therefore, expected performance really depends on the architecture and the compiler.

If the memory bandwidth is large, as for the CRAY Y-MP, we may expect close to peak performance. Smaller bandwidths are almost immediately and inevitably reflected by smaller fractions of the peak performance. It then depends on the compiler how much advantage can be taken from a possible combination of operations in order to reduce the necessity of data transfer.

References

[1] J. J. Dongarra. Performance of various computers using standard linear equations software in a Fortran environment. Computer Science Technical Report CS-89-85, University of Tennessee, March 1990.

[2] J. J. Dongarra, I. S. Duff, D. C. Sorensen, and H. A. van der Vorst. *Solving Linear Systems on Vector and Shared Memory Computers.* SIAM, Philadelphia, 1991.

[3] G. Radicati di Brozolo and M. Vitaletti. *Sparse matrix-vector product and storage representations on the IBM 3090 with Vector Facility.* Technical Report 513-4098, IBM-ECSEC, Rome, July 1986.

Table 1: Speed in Mflops for 50 Iterations of the Iterative Techniques

Machine	Standard ICCG Mflops	Diag-order ICCG Mflops	Scaled CG Mflops	Peak Performance Mflops
NEC SX-3/22 (2.9 ns, 1 proc)	60.4	607	1124	2750
CRAY Y-MP C90 (4.2 ns, 1 proc)	56.0	444	737	952
IBM 9000 Model 820 (1 proc)	43.5	39.6	74.6	444
CRAY Y-MP/464 (6 ns, 1 proc)	38.0	175	261	333
CRAY 2 (4.1 ns, 1 proc)	27.5	96.0	149	500
IBM 3090 Model S (1 proc)	19.1	13.5	32.3	133
IBM RS/6000-550 (24 ns)	18.8	18.3	21.1	81
IBM 9121 (15 ns) (1 proc)	14.1	10.6	25.4	133
DEC Vax/9000 (16 ns, 1 proc)	10.2	9.48	17.1	125
CONVEX C3210 (1 proc)	6.65	15.8	19.1	50
CONVEX C-240 (1 proc)	6.45	11.7	13.8	50
Alliant FX/2800 (1 proc)	2.66	2.18	2.98	40
DEC Vaxstation 5000/200	2.57	2.24	2.84	
DEC Vaxstation 5000/125	1.49	1.33	1.58	

For unpreconditioned CG the matrix A has been scaled so that `A( i, 1 ) = 1` for all i. This property has been exploited in the computation of the matrix-vector product. The incomplete Cholesky preconditioning matrix K (for preconditioned CG) can be written as

$$K = (L + \tilde{D})\tilde{D}^{-1}(\tilde{D} + L^T),$$

in which L^T is the strictly upper triangular part of the matrix A. In this case the matrix A has been scaled symmetrically in such a way that $\tilde{D} = I$. This is, of course, exploited in the preconditioning part. In this case, the $a(i, 1)$ are not necessarily equal to 1.

The block structure of A has not been exploited in our code; the block structure of K has been exploited in the "diagonalwise ordering" variant. The parts of the code for the computation of the matrix-vector product (along with a required inner-product for CG) and the preconditioning part are given in the Appendix.

The information in the tables was compiled over a period of several months. Subsequent systems software and hardware changes may alter the timings to some extent.

One further note: In multiprogramming environments it is often difficult to reliably measure the execution time of a single program. Anyone evaluating machines and operating systems should gather additional data.

The column labeled "Standard ICCG" refers to the use of the standard lexicographical ordering of the unknowns. This approach leads to recurrence relations. These relations can be partly vectorized by loop splitting. This has not been coded explicitly in the body of the preconditioner (see the Appendix routine `prec3s`), but we have included an example in this document (Appendix routine `prec3p`).

The column labeled "Diag-order ICCG" refers to the use of diagonal ordering of the unknowns over the x, y planes of the grid [5], [2]. The compiler should be instructed that there are no recurrences in this case (by means of compiler directives; see the Appendix routine `prec3d`). Diagonal ordering leads to vector code without indirect addressing, but with only modest vector lengths (about 50, on the average). Longer vector lengths can be obtained by the so-called hyperplane ordering (see [2]).

4.1 Obtaining the Software

The software used to generate the data for this report can be obtained by sending electronic mail to *netlib@ornl.gov* .

To receive the single-precision software for this benchmark, in the mail message to *netlib@ornl.gov* type *send sparses from benchmark* .

To receive the double-precision software for this benchmark, type *send sparsed from benchmark* .

will be required for convergence, the overall solution time may be longer.

3.2 Other Preconditioners

A variety of preconditioners are suggested in the literature. The Mflops rates for these preconditioners may vary widely (as well as their effect on the number of iteration steps). Some of these preconditioners are specifically designed for vector computers (e.g., those based on truncated Neumann series) or for parallel computers (e.g., the preconditioners constructed on overlapping domains). Also, some incomplete block decompositions lend themselves more to vector computers, when applied to matrices with regular sparsity patterns. The success of these block decompositions seems more apparent for 2D problems; for 3D problems their effect seems to be less significant with respect to point incomplete decompositions.

Another approach is to apply the preconditioners in combination with some coloring scheme, for example, red-black ordering. While this may lead to higher Mflops rates, it often also leads to significant increases in the number of iteration steps, so that CPU time or wall clock time may not decrease. For more complicated coloring schemes (with up to 20–100 colors), the number of iteration steps is reported to increase only marginally at most.

Finally, one may apply preconditioning with a polynomial in A itself. Needless to say, in this case the Mflops rate of the preconditioned iterative scheme is largely determined by the Mflops rate for the matrix-vector product.

4 Results

In Table 1 we show the results for the various machines. The linear system is generated by discretizing a three-dimensional, second-order, elliptic partial differential equation by finite differences over a rectangular equidistant mesh, with Dirichlet boundary conditions. We do not specify the PDE here, since the performance of the method should not depend on knowledge of the PDE: one is not allowed to take any advantage of the PDE other than the structure of the resulting system.

The four diagonals of the upper triangular part of the symmetric matrix that contain nonzero elements are stored in the first four columns of an array `A( *, 10 )`
`A( i, 1 )` the element on the main diagonal (on the i-th row),
`A( i, 2 )` the element on the first codiagonal (on the i-th row),
`A( i, 3 )` the element on the $(nx+1)$-th codiagonal (i-th row, $nx = 100$),
`A( i, 4 )` the element on the $(nxny+1)$-th codiagonal (i-th row, $nxny = 10,000$).
The 5-th column of `A` is used to store the right-hand side of the linear system, the 6-th column contains the diagonal of the preconditioner, columns 7 through 9 are used as working space for CG, and the 10-th column contains scaling information for the matrix A (this is necessary for backscaling the solution).

Admittedly, the generated problem might be solved more efficiently by other methods (see [4] for a survey of methods). We have tried to avoid comparing iterative schemes. In fact, we are not interested even in the solution of the system. We ignore properties of the given system, other than its symmetry and its nonzero diagonal structure. Our objective is to obtain a quick impression of the performance of a generic sparse matrix solver (see the Appendix for our examples).

We have also avoided discussion of suitable stopping criteria. Thus, we simply carry out 50 iteration steps of the CG method (we have implemented an inexpensive test to avoid excessive iteration steps if one solves smaller systems). These 50 iteration steps allow for a sufficiently accurate CPU timing, and they also enable some realistic optimization of the algorithm—for example, by grouping iteration steps in order to reduce data transfer (see [2]). Moreover, this number is representative of many actual situations and is not too large that running the code is unnecessarily expensive.

3.1 Conjugate Gradient Schemes

We report performance rates for three different solution schemes: scaled CG (i.e., with diagonal scaling of the matrix), standard ICCG, and ICCG with diagonal ordering. Of these three solution schemes scaled CG will often require the most iterations to converge. On the other hand, it is easily vectorizable and parallelizable, and it is cheaper in terms of floating-point operations per iteration step.

Since CG is usually combined with some kind of preconditioning, we have also carried out 50 iteration steps of CG with standard incomplete Cholesky factorization as a preconditioner. The CPU time for the construction of that preconditioner has not been included in our timing.

The performances for these (preconditioned) CG algorithms are thought to be representative of a class of well-structured problem. For less well-structured problems, the performance for scaled CG is more or less determined by the performance for the sparse matrix-vector product (usually determined by indirect addressing). In such cases, the standard lexicographical ordering preconditioned CG gives a reasonable impression of the performance of preconditioned CG. Diagonal ordering is not feasible for more general nonzero structures (though in some cases some improvement can be obtained by reordering the unknowns; see Radicati et al. [3]).

The number of operations for the scaled CG method is about $22n$ flops per iteration for our model problem, whereas the preconditioned versions take $35n$ flops per iteration. (We count both additions and multiplications as distinct floating-point operations.)

In judging the performance of the algorithms, it is important to take into account not only the execution rate for the methods, but also the number of operations per iteration and the number of iterations required to converge. As an example, in Table 1 we see that CG with diagonal scaling has the highest execution rate, however, since more iterations usually

the CG method iterates on a well-structured 7-diagonal model problem:

$$A = \begin{pmatrix}
x & x & & & & x & & & & & & x & & & & \\
x & x & x & & & & x & & & & & & x & & & \\
 & x & x & x & & & & x & & & & & & x & & \\
 & & x & x & x & & & & x & & & & & & x & \\
 & & & x & x & & & & & x & & & & & & \ddots \\
x & & & & & x & x & & & & x & & & & & \\
 & x & & & & \ddots & \ddots & \ddots & & & & \ddots & & & & & \\
 & & \ddots & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & & & \\
x & & & & & & & & & & & & & & & & \\
 & x & & & & & & & & & & & & & & & \\
 & & \ddots & & & & & & & & & & & & & &
\end{pmatrix}.$$

There is even more structure in the matrix A. The matrix A can be seen as a symmetric block tridiagonal matrix:

$$A = \begin{pmatrix}
A_1 & D_1 & & \\
D_1 & A_2 & D_2 & \\
 & D_2 & \ddots & \ddots \\
 & & \ddots &
\end{pmatrix},$$

in which the D_i are diagonal matrices, and the A_i are again symmetric block tridiagonal matrices

$$A_i = \begin{pmatrix}
A_{i,1} & C_{i,1} & & \\
C_{i,1} & A_{i,2} & C_{i,2} & \\
 & C_{i,2} & \ddots & \ddots \\
 & & \ddots &
\end{pmatrix}.$$

The $C_{i,j}$ are diagonal matrices, and the $A_{i,j}$ are symmetric tridiagonal matrices.

2.2 Performance of Other Iterative Solvers

The Mflops rate per iteration step of CG is almost entirely determined by the performance of the matrix-vector product (Ap_i) and the preconditioner (solve w_{i+1} from $Kw_{i+1} = r_{i+1}$). In biconjugate-gradient (Bi-CG) methods, the computational work per iteration step is almost double the work for conjugate-gradient (CG) methods; each pass of the iteration loop now involves a matrix-vector product with A and A^T (and the preconditioned version involves solve steps with K and K^T). In the case that the steps with transpose operators can be computed with similar Mflops rates as for the original operators, one may expect the same Mflops rate for Bi-CG as for CG. In many situations, however, one has a data structure which facilitates the efficient computation of products like Ap_i (e.g., row-wise compressed), and then a product like $A^T\hat{p}_i$ often requires indirect addressing. This, of course, may lead to a reduced Mflops rate for Bi-CG in comparison with CG.

For the quasi-minimum residual (QMR) alorithm[4], in which the bulk of the computational work is the same as for Bi-CG, a similar situation holds. In other variants on Bi-CG the computation of $A^T\hat{p}_i$ has been avoided and has been replaced by a product involving A (also, the solve with K^T has been replaced by a solve with K). The more well known of these schemes are CGS, Bi-CGSTAB, and TFQMR. These schemes lead to almost identical Mflops rates as for CG.

For iterative schemes that are based on the generation of a full orthogonal basis for the search space (e.g., GMRES, ORTHODIR, ORTHORES, ORTHOMIN, and GENCG), the situation is more complicated. Typically, in the full versions (i.e., not truncated or restarted) of these schemes, we see that in the i-th iteration step one matrix-vector product with A, one solve with K (in the preconditioned versions), i inner-products, and i vector updates are involved. For low values of i we therefore may expect similar Mflops rates as for CG, while for increasing values of i the work in the inner-products and the updates tends to dominate, and hence the Mflops rate is increasingly determined by these operations.

3 Test Problem and Conjugate Gradient Solver

For this benchmark, the test problem is generated by discretizing a three-dimensional elliptic partial differential equation by the standard 7-point central difference scheme over a three-dimensional rectangular grid, so that we have 100 unknowns in each direction. This procedure results in a system of order 1,000,000. The system is scaled such that it has a unit diagonal (which is common practice for many iterative schemes). For our test problem,

$p_{-1} = 0; \beta_{-1} = 0$
$\rho_0 = (r_0, r_0)$
for $i = 0, 1, 2, ...$
$p_i = r_i + \beta_{i-1} p_{i-1}$
$q_i = A p_i$
$\alpha_i = \frac{\rho_i}{(p_i, q_i)}$
$x_{i+1} = x_i + \alpha_i p_i$
$r_{i+1} = r_i - \alpha_i q_i$
if x_{i+1} accurate enough then quit
$\rho_{i+1} = (r_{i+1}, r_{i+1})$
$\beta_i = \frac{\rho_{i+1}}{\rho_i}$
end

In this algorithm the variables α, β, and ρ are scalars, x, r, p, and q are vectors of length n, $(x, y) = x^T y$, and A is a sparse n by n symmetric positive definite matrix.

The speed of convergence of the conjugate gradient method depends strongly on the spectrum of the matrix in the system. It is often advantageous to select a matrix, K, as a preconditioner for the system that clusters the eigenvalue of the system. The following computational scheme is for preconditioned CG, for the solution of $Ax = b$ with preconditioner K^{-1}.

x_0= initial guess; $r_0 = b - A x_0$
$p_{-1} = 0; \beta_{-1} = 0$
Solve w_0 from $K w_0 = r_0$
$\rho_0 = (r_0, w_0)$
for $i = 0, 1, 2, ...$
$p_i = w_i + \beta_{i-1} p_{i-1}$
$q_i = A p_i$
$\alpha_i = \frac{\rho_i}{(p_i, q_i)}$
$x_{i+1} = x_i + \alpha_i p_i$
$r_{i+1} = r_i - \alpha_i q_i$
if x_{i+1} accurate enough then quit
Solve w_{i+1} from $K w_{i+1} = r_{i+1}$
$\rho_{i+1} = (r_{i+1}, w_{i+1})$
$\beta_i = \frac{\rho_{i+1}}{\rho_i}$
end

If one transforms the system explicitly to the symmetric system with matrix $L^{-1} A L^{-T} y = L^{-1} b$, then the solution needs to be backtransformed to $x = L^{-T} y$ afterwards. This popular formulation has the advantage that the preconditioner need not be split into two factors. Further, it avoids backtransform solutions and residuals, as is necessary when one applies classical CG to $L^{-1} A L^{-T} y = L^{-1} b$ (with $K = L L^T$).

continually refining an initial approximate solution so that it becomes closer and closer to the correct solution. With an iterative method a sequence of approximate solutions $\{x^{(k)}\}$ is constructed which essentially involve the matrix A only in the context of matrix-vector multiplication. Thus the sparsity can be exploited so that each iteration requires $O(n)$ operations.

As a result of the algorithm, sparsity of the matrix, and the data structures, the time-consuming elements of iterative methods can be quite different from the time-consuming elements of a direct solver. Thus, the performance of an iterative method can be completely different from the performance of a direct solver.

The benchmark program presented here is a first attempt to construct a representative test problem for sparse matrices on a variety of machines. The heart of the code is a conjugate gradient (CG) iterative solver for large, sparse, symmetric and positive definite linear systems. The performance range and difficulties encountered on a specific architecture, made visible by this simplified model problem, are more or less representative for a class of iterative method (Lanczos, Bi-CG, CGS, etc.; see [2] for further details).

The purpose of this report is to describe the performance of various computers when executing a Fortran implementation of common algorithms used in solving sparse matrix problems.

We are not interested in comparing iterative schemes, or in measuring the number of iterations needed to converge, or even in obtaining the solution. Our intent here is to is to provide a means for quickly testing and comparing the performance of a computer on commonly occurring solution methods for solving sparse matrix problems.

In Section 2 we provide an overview of the conjugate gradient algorithm and give pointers to related approaches. Section 3 describes the test problem, the structure of the matrix used in this evaluation, and the preconditioners used in the algorithm. Section 4 describes the tests and displays the performance for the various machines and the approaches used. In addition, instructions for running the test are provided, as well as a description on how the software can be obtained. Section 5 presents some conclusions.

2 Conjugate Gradient Algorithm and Related Algorithms

We begin with a discussion of the classical conjugate gradient algorithm.

2.1 Classical Conjugate Gradient Algorithm

The classical conjugate gradient algorithm for the solution of $Ax = b$ can be represented by the following scheme:

x_0= initial guess; $r_0 = b - Ax_0$

Performance of Various Computers Using Standard Techniques for Solving Sparse Linear Equations

Jack J. Dongarra[‡] Henk A. Van der Vorst[§]

1 Introduction and Motivation

The LINPACK benchmark[1] has become popular in the past few years as a means of measuring floating-point performance on computers. The benchmark shows in a simple and direct way the performance to be expected for a range of machines when doing dense matrix computations.

Many people have discovered that the situation for *sparse* matrices is less clear. For very large, structured sparse systems, if one applies a dense direct solver operating on a banded matrix, one may expect performance similar to that of the dense case. In other situations, however, the performance strongly depends on the nonzero structure of the matrix, the storage scheme used, and the chosen algorithm.

The algorithm used in the LINPACK benchmark is a direct method. That is, the original matrix is factored into the product of two simpler matrices, a lower triangular matrix L and a upper triangular matrix U, and these matrices are used to construct the solution. The factorization requires $O(n^3)$ operations and the solution $O(n^2)$. Thus, the solution can be determined in a fixed number of operations. A direct method can be impractical, however, if the matrix is large and sparse, because the sought-after factors can be dense and the time to construct the factors large.

Many important practical problems give rise to large sparse systems of linear equations. One reason for the great interest in sparse linear equations solvers and iterative methods is the importance of being able to obtain numerical solutions to partial differential equations. Such systems appear in studies of electrical networks; models of economic systems; and physical processes such as diffusion, radiation, and elasticity. Iterative methods work by

*This work was supported in part by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S Department. of Energy, under Contract DE-AC05-84OR21400, and in part by the Stichting Nationale Computer Faciliteit (NCF), the Netherlands

†This version is dated May 26, 1992

‡University of Tennessee and Oak Ridge National Laboratory

§University of Utrecht