DOE/ER/25221--1

GRANT NUMBER DE-FG03-94ER25221

1994-1996 FINAL REPORT

Project Title: Iterative Methods for Large Scale Nonlinear and Linear Systems

Principal Investigator: Homer F. Walker

Address: Mathematics and Statistics Department
Utah State University
Logan, UT 84322-3900

Telephone: (801) 797-2026
Electronic mail: walker@math.usu.edu
FAX: (801) 797-1822

1. Introduction and overview.

The major goal of this research has been to develop improved numerical methods for the solution of large-scale systems of linear and nonlinear equations, such as occur almost ubiquitously in the computational modeling of physical phenomena. A concomitant goal has been to apply these methods to the modeling of complex physical phenomena on high performance computers. Particular target applications have included modeling of reacting flows (combustion and other chemically reacting flows) and other problems in computational fluid dynamics.

The numerical methods of central interest have been Krylov subspace methods for linear systems, which have enjoyed great success in many large-scale applications, and Newton-Krylov methods for nonlinear problems, which use Krylov subspace methods to solve approximately the linear systems that characterize Newton steps. Krylov subspace methods have undergone a remarkable development over the last decade or so and are now very widely used for the iterative solution of large-scale linear systems, particularly those that arise in the discretization of partial differential equations (PDEs) that occur in computational modeling; see [SR9] and [SR10] for surveys and for references to the specific methods mentioned here. Newton–Krylov methods have enjoyed parallel success and are currently used in many nonlinear applications of great scientific and industrial importance. In addition to their effectiveness on important problems, Newton–Krylov methods also offer a nonlinear "framework" within which to transfer to the nonlinear setting any advances in Krylov subspace methods or preconditioning techniques, or new algorithms that exploit advanced machine architectures.

This research has resulted in a number of improved Krylov and Newton–Krylov algorithms together with applications of these to important linear and nonlinear problems. A particular outcome has been the development of a robust and efficient Newton-Krylov solver, implemented in a Fortran code called NITSOL. Another has been the implementation of the algorithm underlying NITSOL as the core nonlinear solver in the code *MPSalsa*, developed in collaboration with researchers at Sandia National Laboratories (SNL) for modelling chemically reacting flows on massively parallel machines. Other notable collaborations have involved researchers at the Center for High Performance Computing at the University of Utah, the Center for Applied Scientific Computing (CASC) at Lawrence Livermore National Laboratory (LLNL) , and the Center for Research in Parallel Computing at Rice University.

Referencing conventions are as follows: Publications acknowledging support from this grant are referenced by [P...] and listed in §2.4. Supplementary references are referenced by [SR...] and listed in §2.5. Invited talks are referenced as [T...] and listed in §2.3.2.

2. Technical accomplishments and activities during the 1994-96 period.

We first outline research on Newton–Krylov methods and then review work on Krylov subspace methods.

2.1. Newton–Krylov methods.

In spite of their conceptual simplicity, Newton–Krylov methods involve complex issues, particularly the following:

- How accurate should each approximate Newton step be, i.e., when should the Krylov iterations be stopped, in order to enhance efficiency and robustness away from a solution and ensure fast convergence near a solution?
- How can the method be "globalized" compatibly with the strategy for stopping the linear iterations and other algorithmic features?
- Which Krylov solver is best suited for the overall method and perhaps a particular problem or computer architecture as well?

A preliminary goal of this research has been to contribute to general, theoretically sound frameworks for Newton–Krylov methods. An ideal context for this is that of an inexact Newton method [SR6].

2.1.1. Inexact Newton methods.

The characteristic property of an inexact Newton method is that each step is required to reduce the norm of the local linear model of the nonlinear residual function F. Specifically, at the kth iteration, the step s_k is required to satisfy an *inexact Newton condition* $\|F(x_k) + F'(x_k)\, s_k\| \leq \eta_k \|F(x_k)\|$ for some $\eta_k \in [0, 1)$. A Newton–Krylov method is naturally a special case: At the kth step, we first choose η_k and then apply the Krylov solver to $F'(x_k)\, s = -F(x_k)$ until the inexact Newton condition holds. Used in this way, the η_k's are often called *forcing terms*. Note that the issue of when to stop the linear iterations is now the issue of choosing η_k. It is shown in [SR6] that local convergence to a solution is controlled by the choices of the η_k's; furthermore, these choices are critical to the efficiency of the method away from a solution and often its robustness as well (see [P2]).

In previous work in [SR7], we have formulated a number of inexact Newton methods with strong global convergence properties. These include a number of methods that are well-suited for practical implementation as Newton–Krylov methods, as well as several abstract methods that can be used to obtain global convergence results for well-known Newton-like methods such as trust region methods.

In follow-up work with S. Eisenstat in [P2], we have addressed the issue of making refined choices of the η_k's. Choices that had been previously proposed result in fast convergence near a solution but do not adequately address efficiency and robustness away from a solution. The central issue is that of *oversolving* the linear system $F'(x_k)\, s = -F(x_k)$, i.e., going to the expense of reducing the linear residual norm, which is also the local linear model norm, without achieving a commensurate reduction in the norm of F itself. In [P2], we offer several very promising choices of the η_k's that tend to minimize oversolving, retain fast convergence near a solution, and, by maintaining good agreement between F and its local linear model, also enhance the robustness of the method.

2.1.2. Newton–Krylov implementations.

A general inexact Newton backtracking method is formulated in [SR7] that has very attractive global convergence properties. In collaboration with M. Pernice at the Utah Center for High Performance Computing, University of Utah, we have developed a Newton–Krylov algorithm based on this method, implemented in a Fortran solver called NITSOL, that allows the use of the sophisticated forcing term choices in [P2], together with several "transpose-free" Krylov solvers and a number of refinements, such as options for evaluating Jacobian-vector products through either user-supplied analytic evaluation routines or through finite-differences of function values. The latter feature allows optional higher-order differencing, including low-cost selective higher-order differencing as in [SR18] with restarted GMRES [SR15]. With NITSOL, users have great flexibility to address challenging problem features through sophisticated preconditioning strategies. In addition, a user-supplied inner-product is allowed; this can be advantageous in treating problem scaling and also allows easy parallelization through a distributed inner-product routine, with no internal modifications of NITSOL itself. A description of NITSOL and experiments demonstrating its features and performance on a set of realistic test problems are written up in [P3]. A precursor algorithm and experiments involving overlapping Schwarz preconditioning on distributed memory machines are described in [P4].

In collaboration with R. S. Tuminaro and J. N. Shadid at SNL, the inexact Newton backtracking algorithm from [SR7] together with forcing term options from [P2] and other features from NITSOL have been implemented as the core nonlinear solver in the parallel reactive flow code *MP-Salsa*. Extensive experiments with this code on a number of benchmark CFD problems and realistic 3D simulations are described in [P5]. These experiments demonstrate the general effectiveness of this globalized inexact Newton approach to the fully-coupled solution of steady-state Navier–Stokes problems, and the resulting code should find wide application at SNL and elsewhere.

2.2. Krylov subspace methods.

Among current Krylov subspace methods, GMRES is the most widely used and is likely to remain of central importance. However, the increasing cost per iteration of GMRES caused much attention to turn in the late 1980's toward short-recursion Lanczos-based methods such as the biconjugate gradient (BCG) and conjugate-gradient squared (CGS) methods. Although BCG and CGS have had some successes where GMRES has not done well, a major shortcoming of these methods is that the residual norms of the iterates often oscillate wildly, which is at least unnerving and can create serious numerical problems as well. This shortcoming has provided important impetus for the development of various BiCGSTAB- and QMR-type methods; these produce residual norm sequences that, although not monotone decreasing, are usually but not always fairly well-behaved. Our work to date has centered around GMRES, "residual smoothing" techniques for general iterative linear solvers, and recent adaptations of Krylov subspace methods to path following problems.

2.2.1. GMRES work.

Our GMRES work was directed toward developing more accurate and more efficient implementations. In joint work with L. Zhou [P8], then a Ph.D. student at Utah State University, we have outlined Gram–Schmidt and Householder implementations of GMRES that are simpler than the usual ones (see [SR15], [SR19], [SR20]). In these, the GMRES least-squares problem emerges in upper-triangular form rather than upper-Hessenberg form, so no Givens rotations are necessary; in addition, the residual vector is always available at no cost in the Gram–Schmidt implementation.

We also worked with P. N. Brown in CASC at LLNL to investigate the performance of GMRES on singular or ill-conditioned systems. This work has resulted in improved understanding of the

behavior of GMRES and other residual minimizing methods on such systems and in implementations of GMRES with improved capabilities for detecting and handling ill-conditioning when it threatens to degrade the performance of the method. This work is written up in [P1].

2.2.2. Residual smoothing techniques.

In previously supported work with L. Zhou [SR26], we investigated "residual smoothing" techniques for general iterative linear solvers. These produce from a given sequence of iterates $\{x_k\}$ an auxiliary sequence $\{y_k\}$ by $y_0 = x_0$ and $y_k = (1 - \xi_k)y_{k-1} + \xi_k x_k$, $k = 1, 2, \ldots$, where the ξ_k's are chosen to produce "smoothly decreasing" residual norm sequences. In [SR26], the focus is on numerically sound implementations of smoothing and on ways in which it can be used to relate various QMR-type methods to underlying Lanczos-based methods such as BCG, CGS, and BiCGSTAB. These results offer valuable practical tools and, in addition, give important new perspectives on QMR-type methods.

In more recent work, we have shown how residual smoothing techniques can be used to extend results of Brown [SR1] and Cullum and Greenbaum [SR3], [SR4] correlating "peaks" and "plateaus" of residual norm sequences produced by certain well-known pairs of Krylov subspace methods. This is written up in [P6].

2.2.3. Krylov subspace methods for path following problems.

In work near the end of the 1994-96 funding period, we began exploring a certain procedure for adapting Krylov subspace methods to solving the underdetermined linear systems that arise in path following problems, including continuation and homotopy problems. This procedure, which can be very economically implemented using Householder transformations, has the advantages of satisfying constraints exactly, regardless of how accurately the underdetermined system is solved, and of not worsening conditioning through inappropriate scaling. This procedure is presented in [P7], which shows promising experimental results on PDE test problems.

During this time, we also contributed to a project with L. T. Watson and others aimed at upgrading the HOMPACK code for continuation and homotopy problems (see [SR21]) to include new Krylov solver options and to exploit features of Fortran 90. The upgraded code is described in [P9], which will appear in the ACM *Transactions on Mathematical Software.*

2.3. Other activities.

2.3.1. Collaborative and advisory activities.

We have noted some collaborations above with P. N. Brown (CASC, LLNL), S. Eisenstat (Computer Science, Yale), M. Pernice (Utah Center for High Performance Computing, University of Utah), R. S. Tuminaro and J. N. Shadid (SNL), L. Zhou (then a graduate student at Utah State University), and with L. T. Watson (Computer Science,Virginia Polytechnic Institute and State University) and other co-authors in [P9]. In addition, we have engaged in other collaborative or advisory activities relating to the work supported by this grant as follows:

- *Ongoing collaboration with members of CASC at LLNL.* The main current effort is with P. N. Brown, A. Hindmarsh, and others revise the Brown–Saad Newton–Krylov code NKSOL [SR2] to include many features of NITSOL. We have also begun discussing with C. Woodward (née San Soucie) the use of the revised code and related methods on time-dependent reservoir simulation problems and other applications of interest to CASC. The usefulness in this context of the backtracking algorithm from [SR7] with forcing terms from [P2] has been demonstrated in her previous work in [SR5] and [SR16].

- *Visits to SNL.* These were a five-week visit in the summer of 1995, followed by a nine-week visit in the summer of 1996. These visits enabled the collaboration with Tuminaro and Shadid on the *MPSalsa* work noted above and made possible extensive experimentation with the modified code on massively parallel machines housed at SNL.
- *Ongoing work with M. Pernice, P. Smith, and others associated with the Utah Center for High Performance Computing at the University of Utah.* The object is to further explore parallel, domain decomposition-based implementations of Krylov subspace methods, preconditioners, and Newton–Krylov methods. Particular target applications are to CFD and reactive flow problems.
- *Work with graduate students and postdoctoral associates.* There were a number of graduate student- and postdoc-related activities not funded by this grant but closely related to the supported research. After receiving his Ph.D. at Utah State University, L. Zhou was sponsored as a postdoc by the Utah Center for High Performance Computing at the University of Utah in cost sharing through another grant with this PI at Utah State, where he collaborated in research on Newton–Krylov methods and applications on high-performance computers (see [P4]). S. Yi received a Ph.D. from Utah State in 1996 under the PI's advisement for dissertation work on Krylov subspace methods. Also, while this grant has been in force, the following interacted with the PI during their dissertation or postdoctoral work and subsequently referenced the PI's work in the publications indicated: M. Sosinkina (VPI&SU, [SR17], [P9]), C. San Soucie (Rice, [SR16], [SR5]), H. Klíe (Rice, [SR12], [SR5]), M. Ramé (Rice, [SR12]), J. Zhang (George Washington, [SR22], [SR23], [SR24], [SR25]), L. Vicente (Rice, [SR11]), R. LeHoucq (Rice, [13]), V. Eijkhout (UCLA, [SR8]).
- *The Workshop on Iterative Methods for Large-Scale Nonlinear Systems.* This workshop was organized by the PI and M. Pernice and held on the Utah State campus September 14–16, 1995. No funding was provided through this grant; however, major partial funding was provided through a related DOE grant (DE-FG03-95ER25255), with leveraging through grants from NSF, Utah State, and the Utah Supercomputing Institute (now the Utah Center for High Performance Computing). The aim of the workshop was to bring together researchers working on large-scale applications with numerical specialists of various kinds. Applications addressed included reactive flows (combustion and other chemically reacting flows, tokamak modeling), porous media flows, cardiac modeling, chemical vapor deposition, image restoration, macromolecular modeling, and population dynamics. Numerical areas included Newton iterative (truncated Newton) methods, Krylov subspace methods, domain decomposition and other preconditioning methods, large-scale optimization and optimal control, and parallel implementations and software. The 34 participants included 25 from universities and 9 from national laboratories. Among these were 10 graduate students and 3 postdoctoral associates, with strong representation of women and minorities. See [SR14] or the online proceedings at `http://www.usi.utah.edu/logan_proceedings` for further information.

2.3.2. Invited talks.

Invited talks by the PI from the beginning of funding in August, 1994, to date on research supported by this grant are as follows:

[T1] *A GMRES-backtracking Newton-iterative method*, Session on Algorithms and Applications, 15th International Symposium on Mathematical Programming, Ann Arbor, Michigan, August, 1994.

[T2] *Krylov subspace methods and residual smoothing techniques*, Sandia National Laboratory (Albuquerque), August, 1994.

[T3] *Newton iterative and inexact Newton methods*, Mathematics Department, University of Utah, May, 1995.

[T4] *GMRES on (nearly) singular systems*, Minisymposium on Iterative Methods for Large Linear Systems of Equations, ICIAM 95: Third International Conference on Industrial and Applied Mathematics, Hamburg, Germany, July, 1995.

[T5] *A Newton–Krylov–Schwarz method for PDE applications*, Minisymposium on Industrial Application of Optimization Methods, ICIAM 95: Third International Conference on Industrial and Applied Mathematics, Hamburg, Germany, July, 1995.

[T6] *Adapting Krylov subspace methods to continuation problems*, BYU-UU-USU Seminar on Nonlinear Analysis and PDE, Brigham Young University, January, 1996.

[T7] *An adaption of Krylov subspace methods to path following*, Copper Mountain Conference on Iterative Methods, Copper Mountain Colorado, April, 1996.

[T8] *Adapting Krylov subspace methods to nonlinear problems*, Numerical Analysis Seminar, UCLA, April, 1996.

[T9] *NITSOL: a Newton iterative solver for nonlinear equations*, Minisymposium on Use of Iterative Methods in Optimization and Nonlinear Equations, 5th SIAM Conference on Optimization, Victoria, British Columbia, May, 1996.

[T10] *Experiments with an inexact Newton method on steady-state Navier–Stokes problems*, Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, September, 1996.

[T11] *Krylov subspace methods for linear and nonlinear systems*, Mathematics Department, Weber State University, October, 1996.

[T12] *Iterative methods for large-scale systems*, Mathematical Sciences Department, Worcester Polytechnic Institute, February, 1997.

[T13] *An approach to continuation using Krylov subspace methods*, Conference on Computational Science for the 21st Century, Tours, France, May, 1997.

[T14] *An adaptation of Krylov subspace methods to path following problems*, Special Session on Lanczos-Type Iterative Methods, Third IMACS International Symposium on Iterative Methods in Scientific Computation, Jackson Hole, Wyoming, July, 1997.

[T15] *Experiments with a Newton–Krylov method on steady-state Navier–Stokes problems*, Special Session on Krylov Subspace Methods, Third IMACS International Symposium on Iterative Methods in Scientific Computation, Jackson Hole, Wyoming, July, 1997.

2.4. Publications.

Publications supported through this grant and acknowledging its support are the following:

[P1] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), 37–51.

[P2] S. C. Eisenstat and H. F. Walker, *Choosing the forcing terms in an inexact Newton method*, SIAM J. Sci. Comput., 17 (1996), pp. 16–32.

[P3] M. Pernice and H. F. Walker, *NITSOL: a Newton iterative solver for nonlinear systems*, Utah State Univ. Math. Stat. Dept. Res. Report 5/96/85, May, 1996, and Utah Supercomputing Institute Report No. 53, June, 1996, to appear in the Special Issue on Iterative Methods, SIAM J. Sci. Comput.

[P4] M. Pernice, L. Zhou, and H. F. Walker, *Parallel solution of nonlinear partial differential equations using a globalized Newton–Krylov–Schwarz method*, Utah Supercomputing Institute

Report No. 48 (revised) and Utah State Univ. Math. Stat. Dept. Res. Report 7/95/82, July, 1995.

[P5] J. N. Shadid, R. S. Tuminaro, and H. F. Walker, *An inexact Newton method for fully coupled solution of the Navier–Stokes equations with heat and mass transport*, Sandia National Labs. Report Sand97-0132J, January, 1997, submitted to J. Comput. Physics.

[P6] H. F. Walker, *Residual smoothing and peak/plateau behavior in Krylov subspace methds*, Special Issue on Iterative Methods for Linear Equations, Appl. Numer. Math., 19 (1995), pp. 279–286.

[P7] ——, *An adaptation of Krylov subspace methods to path following problems*, Utah State University Math. Stat. Dept. Res. Report 12/96/88, December, 1996, submitted to SIAM J. Sci. Comput.

[P8] H. F. Walker and L. Zhou, *A simpler GMRES*, J. Numer. Lin. Alg. Appl., 1(1994), pp. 571–581.

[P9] L. T. Watson, M. Sosonkina, R. C. Melville, A. P. Morgan, and H. F. Walker, *HOMPACK90: A suite of FORTRAN90 codes for globally convergent homotopy algorithms*, Virginia Polytechnic Institute and State Univ. Dept. of Computer Science Tech. Report TR-96-11, June, 1996, to appear in ACM Trans. Math. Software.

2.5. Supplementary references.

[SR1] P. N. Brown, *A theoretical comparison of the Arnoldi and* GMRES *algorithms*, SIAM J. Sci. Stat. Comput., 20 (1992), pp. 58-78.

[SR2] P. N. Brown and Y. Saad, *Hybrid Krylov methods for nonlinear systems of equations*, SIAM J. Sci. Stat. Comput., 11 (1990), pp. 450–481.

[SR3] J. K. Cullum and A. Greenbaum, *Residual relationships within three pairs of iterative algorithms for solving* $Ax = b$, Tech. Report RC 18672, IBM T. J. Watson Research Center, Yorktown Heights, New York, January, 1993; to appear in SIAM J. Matrix Anal. Appl.

[SR4] J. K. Cullum and A. Greenbaum, *Peaks, plateaus, numerical instabilities in Galerkin and minimal residual pairs of methods for solving* $Ax = b$, preprint, 1994.

[SR5] C. N. Dawson, H. Klíe, C. A. San Soucie, and M. F. Wheeler, *A parallel, implicit, cell-centered method for two-phase flow with a preconditioned Newton–Krylov solver*, TICAM Report 96-35, Texas Institute for Computational and Applied Mathematics, University of Texas at Austin, August, 1996.

[SR6] R. Dembo, S. C. Eisenstat, and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), pp. 400–408.

[SR7] S. C. Eisenstat and H. F. Walker, *Globally convergent inexact Newton methods*, SIAM J. Optimization, 4 (1994), pp. 393–422.

[SR8] V. Eijkhout, *On factorizations of the Hessenberg matrics arising from polynomial iterative methods*, CAM Report 96-44, Dept. of Math., UCLA, October, 1996.

[SR9] R. W. Freund, G. H. Golub, and N. M. Nachtigal, *Recent advances in Lanczos-based iterative methods for nonsymmetric linear systems*, Acta Numerica 1992, Cambridge University Press, 1992, pp. 57–100.

[SR10] M. H. Gutknecht, *Lanczos-type solvers for nonsymmetric linear systems of equations*, Acta Numerica 1997, Cambridge University Press, 1997, pp. 271–397.

[SR11] M. Heinkenschloss and L. N. Vicente, *Analysis of inexact trust-region interior-point SQP algorithms*, Rice Univ. Comput. Appl. Math. Dept. Tech. Report TR95-18, June, 1995.

[SR12] H. Klíe, M. Ramé, and M. F. Wheeler, *Krylov-secant methods for solving systems of nonlinear equations*, Rice Univ. Comput. Appl. Math. Dept. Tech. Report TR95-27, Sept., 1995, submitted for publication.

[SR13] R. B. LeHoucq, *Analysis and Implementation of an Implicitly Restarted Arnoldi Iteration*, Ph.D. thesis, Computational and Applied Mathematics Department, Rice University, 1995.

[SR14] M. Pernice and H. F. Walker, *The Workshop on Iterative Methods for Large-Scale Nonlinear Systems*, Utah State Univ. Math. Stat. Dept. Res. Report 12/95/84, December, 1995.

[SR15] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual method for solving non-symmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[SR16] C. A. San Soucie, *Mixed finite element methods for variable saturated subsurface flow*, Ph.D.thesis, Computational and Applied Mathematics Department, Rice University, 1996.

[SR17] M. Sosonkina, L. T. Watson, R. K. Kapania, *A new adaptive GMRES algorithm for achieving high accuracy*, submitted to SIAM J. Sci. Comput., 1996.

[SR18] K. Turner and H. F. Walker, *Efficient high accuracy solutions with GMRES(m)*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 815–825.

[SR19] H. F. Walker, *Implementation of the GMRES method using Householder transformations*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 152–163.

[SR20] ——, *Implementations of the GMRES method*, Computer Physics Communications, 53 (1989), pp. 311–320.

[SR21] L. T. Watson, S. C. Billups, and A. P. Morgan, *Algorithm 652: HOMPACK: A suite of codes for globally convergent homotopy algorithms*, ACM Trans. Math. Software, 6 (1980), pp. 252–260.

[SR22] J. Zhang, *Accelerated multigrid high accuracy solution of the convection-diffusion equation with high Reynolds numbers*, George Washington University Math. Dept. Report, Dec., 1995, to appear in Numerical Methods for Partial Differential Equations

[SR23] ——. *Minimum residual smoothing in multi-level method*, George Washington University Math. Dept. Report, Jan., 1996, to appear in Applied Mathemathics and Computation.

[SR24] ——, *Analysis of minimal residual smoothing in multigrid*, George Washington University Math. Dept. Report, Feb., 1996.

[SR25] ——, *Multigrid with inexact minimal residual smoothing*, George Washington University Math. Dept. Report, March, 1996.

[SR26] L. Zhou and H. F. Walker, *Residual smoothing techniques for iterative methods*, Special Section on Iterative Methods in Numerical Linear Algebra, SIAM J. Sci. Comput., 15 (1994), pp. 297–312.

DISCLAIMER

DISCLAIMER