SIAM REVIEW
Vol. 31, No. 1, pp. 50–102, March 1989

003

SOME HISTORY OF THE CONJUGATE GRADIENT AND LANCZOS ALGORITHMS: 1948-1976*

GENE H. GOLUB† AND DIANNE P. O'LEARY‡

Abstract. This paper gives some of the history of the conjugate gradient and Lanczos algorithms and an annotated bibliography for the period 1948-1976.

Key words. conjugate gradient algorithm, Lanczos algorithm, variable metric algorithms

AMS(MOS) subject classifications. 65F10, 65F15, 65H10, 65F03

1. Introduction. The conjugate gradient and Lanczos algorithms for solving linear systems of equations and eigenproblems represent very important computational innovations of the early 1950s. These methods came into wide use only in the mid-1970s. Shortly thereafter, vector computers and massive computer memories made it possible to use these methods to solve problems which could not be solved in any other way. Since that time, the algorithms have been further refined and have become a basic tool for solving a wide variety of problems on a wide variety of computer architectures. The conjugate gradient algorithm has also been extended to solve nonlinear systems of equations and optimization problems, and this has had tremendous impact on the computation of unconstrained and constrained optimization problems.

The purpose of this paper is to trace some of the history of the conjugate gradient and Lanczos algorithms during the period from their original development to their widespread application in the mid-1970s.

It is not the purpose of this paper to give the definitive derivation of the algorithms and their properties; for such information, the reader is referred to the references in the bibliography as well as more recent treatments such as *Matrix Computations* by G. H. Golub and C. F. Van Loan (The Johns Hopkins University Press, Baltimore, Maryland, 1983, Chapters 9 and 10). It is necessary, however, to establish notation to make the differences among the variants of the algorithm more clear. This is our first task.

The conjugate gradient algorithm can be thought of as a method for minimizing a function $\frac{1}{2}(x, Ax) - (x, b)$ where A is an $n \times n$ matrix (or an operator on a Hilbert space) and x and b are vectors in the domain and range spaces, respectively. The minimizer of this function satisfies the equation $Ax = b$ if A is self-adjoint and positive definite, so the algorithm provides a means for solving linear equations. It is characterized by the property of stepping at each iteration in the direction of the component of the gradient A-conjugate to the preceding step direction, and by the property of finite termination under exact arithmetic. The residual at step k can be expressed as the product of an optimal polynomial in A with the initial residual, and thus the distribution of eigenvalues is

* Received by the editors June 8, 1987; accepted for publication (in revised form) April 19, 1988.

† Computer Science Department, Stanford University, Stanford, CA 94305. This author's research was partly supported by National Science Foundation Grant DCR84-12314.

‡ Computer Science Department and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. This author's research was partly supported by Air Force Office of Scientific Research Grant AFOSR-82-0078.

exploited; indeed, the method will converge in k or fewer iterations if the operator has k distinct eigenvalues. For simplicity, we will take the initial guess $x_0 = 0$, giving an initial residual, or negative gradient, of $r_0 = b$, and we take this as our first step direction p_0 as well. The algorithm is

$$x_{k+1} = x_k + \alpha_k p_k,$$

$$r_{k+1} = r_k - \alpha_k A p_k,$$

$$p_{k+1} = r_{k+1} + \beta_k p_k,$$

$$\alpha_k = \frac{(r_k, r_k)}{(p_k, Ap_k)}, \quad \beta_k = \frac{(r_{k+1}, r_{k+1})}{(r_k, r_k)}.$$

Variants of the algorithm arise from using different choices of the inner product, from computing the residual directly from its definition as

$$r_{k+1} = b - Ax_{k+1},$$

and from using different, but mathematically equivalent, formulas for the parameters, such as

$$\alpha_k = \frac{(r_k, p_k)}{(p_k, Ap_k)}, \quad \beta_k = -\frac{(r_{k+1}, Ap_k)}{(p_k, Ap_k)}.$$

An important variant can be derived by adding a preconditioning operator M to the formulation, applying the algorithm to the equivalent problem of minimizing $\frac{1}{2}(M^{1/2}y, AM^{1/2}y) - (M^{1/2}y, b)$, and then changing back to the variable $x = M^{1/2}y$. From one point of view, conjugate gradients is a convergence acceleration method for an underlying iteration $My_{k+1} = Ny_k + b$ where $A = M - N$, and this has been a very fruitful insight.

Another equivalent version of the algorithm is formed by eliminating the vectors p in the equations above, giving the three-term recurrence relation

$$x_{k+1} = x_{k-1} + \omega_{k+1}(r_k / \rho_k + x_k - x_{k-1}),$$

$$r_{k+1} = r_{k-1} + \omega_{k+1}(-Ar_k / \rho_k + r_k - r_{k-1}),$$

$$\omega_{k+1} = \left[1 - \frac{(r_k, r_k)}{(r_{k-1}, r_{k-1})} \frac{\rho_{k-1}}{\rho_k} \frac{1}{\omega_k}\right]^{-1}, \qquad \rho_k = \frac{(r_k, Ar_k)}{(r_k, r_k)},$$

with the definitions $x_{-1} = x_0$ and $\omega_0 = 0$.

The idea behind the Lanczos algorithm for determining eigenvalues of a matrix can be discussed easily using the three-term recurrence relation above. By forming a matrix R_k whose columns are the first k residual vectors normalized to length 1, we can derive the relation

$$AR_k = R_k T_k + r_k e_k^T,$$

where T_k is a tridiagonal matrix of dimension k whose elements can be computed from the ρ_j and ω_j values, and e_k is the kth unit vector. The residual vectors are mutually orthogonal, so a full n steps of the algorithm yield a similarity transformation of the matrix A into tridiagonal form if the residual does not become zero prematurely. The intermediate matrices T_k have interlacing eigenvalues, however, and even for small k, some eigenvalues of A may be well approximated. Lanczos used the recurrence relations as a convenient way of constructing characteristic polynomials of the matrices T_k, from which approximations to eigenvalues could be computed.

The algorithm can be extended to minimization of nonquadratic functions $f(x)$. In this case we define the residual vector r_k to be the negative gradient of f evaluated at x_k, and the matrix A, the Hessian matrix for f, changes at each iteration. The alternate formulas no longer give equivalent algorithms, and much study has been given to appropriate choices.

Our discussion does not even hint at the richness of the algorithms: acceleration procedures, convergence properties, and the interplay among the complementary views of the quadratic algorithm as a minimization procedure, as a linear system solver, and as a similarity transformation. The remainder of the paper, devoted to the history of this family of algorithms, focuses on discoveries of these ideas and others. In §2, we trace the early developments at the National Bureau of Standards. Some of the key developments involved in making the algorithms practical are summarized in §3. Section 4 gives information on the organization of the annotated bibliography, §5 is devoted to acknowledgments, and the bibliography and author index follow.

2. Early developments. The original development of this family of algorithms was carried out by a core of researchers including Cornelius Lanczos and Magnus Hestenes at the Institute for Numerical Analysis in the National Applied Mathematics Laboratories at the United States National Bureau of Standards (NBS) in Los Angeles, and Eduard Stiefel of the Eidg. Technische Hochschule Zürich. The first communication among the NBS researchers and Stiefel concerning this algorithm seems to have been at a Symposium on Simultaneous Linear Equations and the Determination of Eigenvalues, held at INA in August 1951, as discussed in Stiefel (1952). Further perspective on the interplay between the researchers at the National Bureau of Standards can be found in Forsythe, Hestenes, and Rosser (1951), Hestenes and Stiefel (1952), Rosser (1953), and Todd (1975). The National Bureau of Standards developments can be traced through internal quarterly project reports of the National Applied Mathematics Laboratories. The following is a condensation of some of the information in those reports. In some cases, work can be attributed to a single person, in others, only to a project directed by a manager or group of managers.

2.1 April - June, 1949.

Project: Determination of Characteristic Values of Matrices. Manager: C. Lanczos.

Lanczos was investigating eigenvalue algorithms in this and other projects and was preparing to write up the work.

Project: Approximate Solution of Sets of Arbitrary Simultaneous Algebraic Equations. Manager: C. Lanczos.

A "method for solving equations was investigated." The method is steepest descent applied to minimizing the residual. "At present the investigation is directed to the possibility of increasing the convergence of the successive reductions, by replacing A by $\bar{A} = \gamma\, 1 + A$ with suitably chosen γ."

This was the quarter in which Hestenes was hired (although in his position at the University of California in Los Angeles since 1947 he had already had contact with the NBS group), but there seems to be no explicit mention of his activities.

2.2 July, 1949 - June, 1951.

Lanczos seems to have been working on the eigenvalue algorithm and other things, and the Lanczos (1950) manuscript was submitted and accepted for publication in the *Journal of Research of the National Bureau of Standards* in the last quarter of 1949. Hestenes seems to have been working on the Hestenes-Karush project and variational problems, among other things. Both were participating in common seminars and project meetings. There is no mention of the conjugate gradient algorithm.

2.3 July - September, 1951.

Project: Solution of Sets of Simultaneous Algebraic Equations and Techniques for the Inversion and Iteration of Matrices. Managers: Forsythe, Hestenes, Lanczos, Motzkin, Rosser, Stein.

"Experimental work with the finite step methods described by M. R. Hestenes in a paper entitled 'Iterative methods for solving linear equations' was initiated by G. E. Forsythe and M. L. Stein." "Dr. E. Stiefel and Dr. M. R. Hestenes are writing a joint paper on extensions and implications of the methods described in the papers presented by J. Barkley Rosser, E. Stiefel and M. R. Hestenes at the Symposium on Simultaneous Linear Equations and the Determination of Eigenvalues held August 23-25, 1951, at the INA. ... For the extensive work of C. Lanczos on the solution of linear algebraic equations, see the description of the algorithm which he devised."

Project: Studies in Applied Mathematics. Managers: Lanczos, Rosser, van der Corput.

This describes the nucleus of his "Solution of systems of linear equations by minimized iterations" paper.

Project: Calculation of Eigenvalues, Eigenvectors, and Eigenfunctions of Linear Operators.

Experiments were conducted on applying Newton's method to the characteristic polynomial obtained from using conjugate gradients ($r - p$ version) on a symmetric positive definite matrix.

2.4 October - December, 1951.

Project: Solution of Sets of Simultaneous Algebraic Equations and Techniques for the Inversion and Iteration of Matrices. Managers: Forsythe, Hestenes, Lanczos, Motzkin, Rosser, Stein.

"The joint exposition by E. Stiefel and M. R. Hestenes of their 'finite iteration' procedure is almost finished."

Lanczos was working on a method to solve $Ax = b$ by finding the large eigenvalues and corresponding eigenvectors of $A' = \lambda_{\max} I - A$, where $\lambda_{\max}$ is the largest eigenvalue of A. He applied a Chebyshev iteration to eliminate components of the residual corresponding to large eigenvalues of A, and resolved components corresponding to small eigenvalues using the eigeninformation for A'. The method was recommended for problems with multiple right-hand-sides.

Project: Variational Methods

Hestenes and Stein completed a study of algorithms for minimizing $(Ax - b)^* H(Ax - b)$. Hayes developed convergence theorems applicable to Rayleigh-Ritz and conjugate gradients for solving linear boundary value problems.

2.5 January - March, 1952.

The Lanczos (1952) manuscript was accepted for publication in the *Journal of Research of the National Bureau of Standards.*

2.6 April - June, 1952.

Project: Solution of Sets of Simultaneous Algebraic Equations and Techniques for the Inversion and Iteration of Matrices Managers: Forsythe, Hestenes, Lanczos, Lehmer, Motzkin.

The Hestenes and Stiefel (1952) manuscript was completed and accepted for publication. "It gives a full description of a wide class of methods of 'conjugate directions,' which includes as special cases both Gaussian elimination and the Hestenes-Lanczos-Stiefel method of 'conjugate gradients.' The latter is a finite iterative scheme which appears practically and theoretically to be a most attractive machine method."

"To test the stability of the conjugate gradient method in regard to rounding-off errors, symmetric systems of linear equations in 12 unknowns were solved on the IBM equipment. In order to know the eigenvalues, an orthogonal matrix was constructed, so that for any given set of eigenvalues a symmetric matrix could be found. The experiments were carried out on three machines [sic] with the following ratios of the largest to the smallest eigenvalue: 4.9, 100, 5000. The computing machine which was used for these experiments had a fixed decimal point and was allowed to work with 10 digits. By shifting, at least seven digits were carried through the computations. For the smallest ratio an answer with seven correct digits was reached in 11 steps. For the ratio 100 six correct digits in the 15th step were obtained. In the third case a good answer has not yet been found since the shifting caused considerable difficulties. The experiments showed the necessity of using floating operations for this method."

Experiments with the nonsymmetric formulas of matrices of dimension 8 gave convergence in less than or equal to 8 steps, even on singular problems, using the SWAC with 8 ½ digits in the arithmetic, obtaining 7-8 correct digits at termination.

Hayes was finishing work on the application of the "method given by E. Stiefel and M. R. Hestenes" to linear boundary value problems. Lanczos was working on solution of large-scale linear systems by Chebyshev polynomials.

2.7 July - September, 1952.

The conjugate gradient algorithm is often called the "Hestenes and Stiefel" algorithm in the reports of numerical experiments and other activities.

3. Key developments related to the algorithms.

3.1 The early papers. The first presentation of conjugate direction algorithms seems to be by Fox, Huskey, and Wilkinson (1948) who considered them as direct methods, and by Forsythe, Hestenes, and Rosser (1951), Hestenes and Stiefel (1952), and Rosser (1953) who discuss the NBS research. The conjugate gradient algorithm was described in Hestenes (1951), Lanczos (1952), Hestenes and Stiefel (1952), Stiefel (1952), Stiefel (1953), Curtiss (1954), Hestenes (1955), Hestenes (1956), and Stiefel (1957). Hestenes, Lanczos, and Stiefel clearly considered the algorithm to be a full n step direct method for certain problems, but also suggested its use as an iterative algorithm requiring fewer than n steps for well-conditioned problems and possibly more than n steps for ill-conditioned ones. Computational results for the conjugate gradient algorithm were presented in Stiefel (1953), Fischbach (1956), Stiefel (1958), and Engeli, Ginsburg, Rutishauser and Stiefel (1959), as well as in Hestenes and Stiefel (1952). Preconditionings, filtering, or change of inner product were considered in Fischbach (1956), Stiefel (1958), and in Engeli et al. (1959). Hayes (1954) and Altman (1959) discuss the conjugate gradient algorithm in Hilbert space.

Computations using the Lanczos algorithm were given in Lanczos (1950) and Rosser, Lanczos, Hestenes, and Karush (1951). Complete reorthogonalization was recommended by Brooker and Sumner (1956), Gregory (1958), and Wilkinson (1958).

3.2 Work in the 1960s. In this period, the conjugate gradient algorithm began to acquire a mixed reputation. It was still regarded as a standard algorithm, as evidenced by its inclusion in the Handbook Series (Ginsberg (1963)), in anthologies such as Beckman (1960), and in review articles in control (Paiewonsky (1965) and Westcott (1969)), chemistry (Wilde (1965)), and pattern recognition (Nagy (1968)).

Frank (1960) tested the algorithm on a matrix with eigenvalues related to the Chebyshev polynomials, the hardest case for conjugate gradients, and reported slow convergence. Applications in structural analysis by Livesley (1960) were unsuccessful, although Bothner-By and Naar-Colin (1962) were satisfied with their results in analysis of chemical spectra, and Feder (1962) recommended the algorithm in lens design. Campbell (1965) studied ocean circulation and Wilson (1966) solved optimal control problems with the aid of conjugate gradients. Pitha and Jones (1967) were other users of the algorithm.

Work was also being done on understanding the s-dimensional steepest descent algorithm, which produces the same sequence of iterates as the conjugate gradient algorithm restarted every s steps. References include Khabaza (1963), Forsythe (1968), and Marchuk and Kuznecov (1968).

Ideas that would eventually lead to successful preconditioned conjugate gradient algorithms were being developed. Dufour (1964) applied conjugate gradients to problems in geodesy and discussed several important ideas, including extensions to least squares problems with equality constraints, preconditioning, and elimination of half of the unknowns using a Schur complement. Varga (1960) suggested a sparse partial factorization of a matrix as a splitting operator for Chebyshev acceleration, and Dupont, Kendall, and Rachford (1968), Dupont (1968), and Stone (1968) also considered sparse factorizations. Other ideas related to preconditioning were discussed by Frank (1960) (polynomial filtering), Wachspress and Habetler (1960) (diagonal scaling), Habetler and Wachspress (1961) (Chebyshev acceleration of SSOR), Ehrlich (1964) (Chebyshev acceleration of block SSOR), Gunn (1964) (Chebyshev acceleration of ADI and ADI on a simpler operator), and Evans (1968) (Chebyshev acceleration of matrix splittings).

Wachspress (1963) used ADI as a preconditioner to conjugate gradients to obtain a very efficient algorithm.

Antosiewicz and Rheinboldt (1962), Nashed (1965), Daniel (1967), Horwitz and Sarachik (1968), Hestenes (1969), and Kawamura and Volz (1969) discussed the conjugate gradient algorithm in Hilbert space, and Kratochvil (1968) studied the algorithm for a class of operators on Banach spaces.

A very important advance in the solution of nonlinear equations and optimization algorithms was made in the development of methods which can solve many such problems effectively without evaluation of the derivative matrix. The first algorithms in this class, which reduce to conjugate gradients on quadratic functions, were presented by Feder (1962), Powell (1962), Fletcher and Powell (1963) building on work of Davidon (1959), Fletcher and Reeves (1964), Shah, Buehler, and Kempthorne (1964), and Broyden (1965). Polak and Ribiere (1969), Polyak (1969), Zoutendijk (1960), Sinnott and Luenberger (1967), Pagurek and Woodside (1968), Luenberger (1969), and Miele, Huang, and Heideman (1969) solved constrained problems using conjugate gradients.

The theory of Kaniel (1966) greatly increased the understanding of the convergence properties of the conjugate gradient and Lanczos methods.

Causey and Gregory (1961), Wilkinson (1962), Wilkinson (1965), and Yamamoto (1968) gave practitioners further insight into causes of failure in the Lanczos algorithm for nonsymmetric problems. The algorithm was used in applications problems in infrared spectra (Eu (1968)), scattering theory (Garibotti and Villani (1969)), network analysis (Marshall (1969)), and nuclear shell analysis (Sebe and Nachamkin (1969)).

3.3 The early 1970s. Although it is clear from the discussion above that the conjugate gradient and Lanczos algorithms were widely used in the 1960s, the numerical analysis community was not satisfied with the understanding of the algorithms or with their speed. Preconditioning techniques were not widely known (although much development had been done on splitting techniques), and it was in the early 1970s that key developments were made in making preconditioned algorithms practical.

The paper of Reid (1971) drew the attention of many researchers to the potential of the algorithm as a iterative method for sparse linear systems. It was a catalyst for much of the subsequent work in conjugate gradients.

The dissertation of Paige (1971), with publications as Paige (1972) and (1976), served the analogous purpose for the Lanczos algorithm by providing, among other things, the first step to an understanding of the loss of orthogonality of the Lanczos vectors, thus giving the key to the development of stable algorithms that did not require complete reorthogonalization. This made the Lanczos algorithm practical for large sparse problems by reducing storage and computation time. Developments along this line were made by Takahasi and Natori (1971-72) and Kahan and Parlett (1976).

Preconditioning techniques, although discussed in the 1960s, now became widely used. Axelsson (1972) suggested preconditioning conjugate gradients by a scaled SSOR operator. Other preconditionings were discussed by Evans (1973), Bartels and Daniel (1974), Chandra, Eisenstat, and Schultz (1975), Axelsson (1976), Concus, Golub, and O'Leary (1976), Douglas and Dupont (1976) and by Meijerink and van der Vorst (partial factorizations) in work that reached journal publication in 1977.

Paige and Saunders (1975) provided the first stable extension of the conjugate gradient algorithm to indefinite matrices. Concus and Golub (1976) considered a class of nonsymmetric matrices.

The block Lanczos algorithm was developed in Cullum and Donath (1974) and Underwood (1975).

Applications of the conjugate gradient algorithm, such as those by De and Davies (1970), Kamoshida, Kani, Sato, and Okada (1970), Kobayashi (1970), Powers (1973), Wang and Treitel (1973), Dodson, Isaacs, and Rollett (1976), and Konnert (1976) and of the Lanczos algorithm, such as those by Chang and Wing (1970), Emilia and Bodvarsson (1970), Weaver and Yoshida (1971), Whitehead (1972), Harms (1974), Hausman, Bloom, and Bender (1975), Ibarra, Vallieres, and Feng (1975), Platzman (1975), Cline, Golub, and Platzman (1976), and Kaplan and Gray (1976), also continued during this period. The Lanczos algorithm was rediscovered by Haydock, Heine, and Kelley (1972) and (1975) and applied to determining energy states of electrons.

3.4 Preconditioning. The *word* "preconditioning" is used by Turing (1948) and by then seems to be standard terminology for problem transformation in order to make solution easier. The first application of the word to the idea of improving the convergence of an iterative method may be by Evans (1968), and Evans (1973) and Axelsson (1974) apply it to the conjugate gradient algorithm. The *idea* of preconditioning the conjugate gradient algorithm is much older than this, as noted above, being perhaps implicit in the original conjugate gradient papers, somewhat more explicit in Hestenes (1956), and actually used in Engeli et al. (1959). Wachspress (1963) seems to be the first to use an iterative algorithm for discretized partial differential equations (ADI) as a preconditioner for the conjugate gradient algorithm.

4. The form of the annotated bibliography. The references are arranged alphabetically by author within year of publication. Each paper is given one or more "Classification Codes":

A applications
C conjugate direction/gradient algorithms for linear systems

E eigenproblems
L Lanczos algorithm for eigenproblems
N nonlinear conjugate gradient algorithms
P preconditioning
S matrix splittings

An author index follows the bibliography.

A reader should keep in mind several warnings. Because of publication delays, alphabetization, and mixing of journal publications with technical reports and dissertations, the bibliography is not completely chronological. The bibliography is not exhaustive; in particular, the references to nonlinear versions of the algorithm represent only a sample of the work done in this period, and references to literature in languages other than English are quite incomplete. The annotation for each paper only gives information relevant to the conjugate gradient and Lanczos algorithms and to preconditioning, and thus may not provide a complete summary of the work.

Quotations in the annotations are excerpts from the work itself. In works concerning applications to partial differential equations, the parameter h denotes the stepsize in the discretization.

5. Acknowledgments. Many people provided encouragement, references, corrections, and additions for this project, and we are very grateful to them. Included in this list are Åke Björck, Paul Concus, James Daniel, Howard Elman, L. Ehrlich, D. J. Evans, Roland Freund, S. Hammarling, B. P. Il'in, B. Kamgar-Parsi, David Kincaid, David Luenberger, Stephen Nash, Chris Paige, Tom Patterson, M. J. D. Powell, Axel Ruhe, Michael Saunders, Paul Saylor, Horst Simon, G. W. Stewart, John Todd, Henk van der Vorst, Eugene Wachspress, Olof Widlund, Ralph Willoughby, H. Woźniakowski, and David Young. The bibliography of Kammerer and Nashed (1972) was also very helpful. The project reports for the National Applied Mathematics Laboratory were obtained through the library of the Center for Applied Mathematics of the National Bureau of Standards (recently renamed the National Institute for Standards and Technology), in Gaithersburg, Maryland.

Annotated Bibliography

1948

1. /C/ Fox, L., H. D. Huskey, and J. H. Wilkinson (1948) "Notes on the Solution of Algebraic Linear Simultaneous Equations," *Quart. J. of Mech. and Appl. Math.* 1, pp. 149-173.

   Presents a "method of orthogonal vectors" as a direct method involving forming an A-conjugate set by Gram-Schmidt orthogonalization and expanding the solution vector in this basis.

2. /P/ Turing, A. M. (1948) "Rounding-off Errors in Matrix Processes," *Quart. J. of Mech. and Appl. Math.* 1, pp. 287-308.

   Introduces a quantitative measure of conditioning. "There is a very large class of problems which naturally give rise to highly ill-conditioned equations [an example being a polynomial fit in two dimensions with data in a small region]. In such a case the equations might be improved by a differencing procedure, but this will not necessarily be the case with all problems. Preconditioning of equations in this way will always require considerable liaison between the experimenter and the computer, and this will limit its applicability" (p. 299).

1950

3. /CEL/ Lanczos, C. (1950) "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators," *J. Res. Nat. Bur. Standards* 45, pp. 255-282.

   Gives a polynomial expansion which can be used to solve the eigenvalue problem and develops recurrence relations for the polynomials. Notes that the recurrence is sensitive to round-off, and develops an alternate one, based on the principle of choosing the combination of previous vectors which makes the norm of the resulting vector as small as possible, achieving a three-term recurrence for the polynomials. Derives a bi-orthogonalization algorithm for finding eigenvalues of nonsymmetric matrices, and derives an algorithm with a single set of vectors for symmetric matrices. Uses the vectors to generate a set of polynomials which accumulate the characteristic polynomial of the original matrix. Recognizes that fewer than n steps may be needed to obtain a subset of the eigenvalues. Presents applications to eigenvalue problems in differential equations and integral operators. "The present investigation contains the results of years of research in the fields of network analysis, flutter problems, vibration of antennas, solution of systems of linear equations, encountered by the author in his consulting and research work for the Boeing Airplane Co., Seattle, Wash. The final conclusions were reached since the author's stay with the Institute for Numerical Analysis, of the National Bureau of Standards." "The literature available to the author showed no evidence that the methods and results of the present investigation have been found before. However, A.M. Ostrowski of the University of Basle and the Institute for Numerical Analysis informed the author that his method parallels the earlier work of some Russian scientists; the references given by Ostrowski are: A. Krylov, *Izv. Akad. Nauk SSSR 7, 491 to 539 (1931);* N. Luzin, *Izv. Akad. Nauk SSSR 7, 903 to 958 (1931).* On the basis of the reviews of these papers in the *Zentralblatt,* the author believes that the two methods coincide only in the point of departure. The author has not, however, read these Russian papers."

4. /EL/ Milne, W. E. (1950) "Numerical Determination of Characteristic Numbers," *J. Res. Nat. Bur. Standards* 45, pp. 245-254.

   Approximates eigensystem of an ordinary differential equation by using a related partial differential equation and discretizing using finite differences. Derives a related trigonometric expansion whose roots determine the eigenvalues of the finite difference system. Relates the method to Lanczos (1950).

1951

5. /EL/ Arnoldi, W. E. (1951) "The Principle of Minimized Iterations in the Solution of the Matrix Eigenvalue Problem," *Quarterly of Appl. Math.* 9, pp. 17-29.

   Derives the nonsymmetric Lanczos algorithm as a Galerkin method with the left and right vectors bi-orthogonal, reducing the matrix to tridiagonal form and proposes its use as an iterative method for n steps or fewer. Derives a new algorithm with the left and right vectors equal and orthogonal, reducing the matrix to upper Hessenberg form. Suggests using several steps of the power method to get a starting vector for either algorithm.

6. /C/ Forsythe, G. E., M. R. Hestenes, and J. B. Rosser (1951) "Iterative Methods for Solving Linear Equations," *Bull. Amer. Math. Soc.* 57, p. 480.

(Abstract for Summer Meeting in Minneapolis, Sept. 4-7, 1951, quoted in its entirety) "Several iterative methods are given for solving the equations $Ax=b$, where A is a given matrix and b is a vector. These methods appear to be particularly adapted to high speed computing machines. They have the property that if there were no round-off error the solution would be obtained in at most n steps where n is the rank of A. In the singular case the least square solution is obtained. At each iteration the problem is started anew. Accordingly there is no accumulation of errors. In the hermitian case the method is based on the following result. If $A,B>0$ are hermitian matrices which commute then the system $b, Ab, \cdots, A^n b$ may be replaced by a set of B-orthogonal vectors by the algorithm $z_0=b$, $z_1=z_0-a_0Az_0$, $z_{i+1}=b_iz_i-a_iAz_i+c_{i-1}z_{i-1}$. (Received July 23, 1951)"

7. /C/ Hestenes, Magnus R. (1951) *Iterative Methods for Solving Linear Equations,* NAML Report 52-9, July 2, 1951, National Bureau of Standards, Los Angeles, California.

(Superceded by Hestenes and Stiefel (1952). Reprinted in *J. of Optimization Theory and Applications* 11 (1973), pp. 323-334.) "The methods presented are an outgrowth of discussions with Forsythe, Lanczos, Paige, Rosser, Stein, and others. For the positive Hermitian case it is convenient to separate the methods into two types. The first [the three-term recurrence form of conjugate gradients] is a method which is my interpretation of the suggestions made by Forsythe and Rosser. The second [the x-r-p version] is one which grew out of my discussion of the problem with Paige. The two methods are equivalent and yield the same estimates at each stage." If the first algorithm is used, recommends three-term recurrence for x with direct calculation of the residual vector. Gives alternate formulas for α and β and relates the parameters in the two methods. Shows finite termination, orthogonality of the residuals, and a bound on α. Discusses the case A positive semi-definite and recommends normal equations for nonsymmetric problems. Relates the algorithm to conjugate direction methods and constructs the formula for the inverse of A. Derives the property that the kth iterate of the algorithm minimizes a quadratic function over a k dimensional subspace. Gives a 4×4 example.

8. /EL/ Hestenes, Magnus R. and William Karush (1951a) "A Method of Gradients for the Calculation of the Characteristic Roots and Vectors of a Real Symmetric Matrix," *J. Res. Nat. Bur. Standards* 47, pp. 45-61.

Uses the iteration $x_{k+1}=x_k-\alpha p(x_k)$, with $p(x)=Ax-\mu(x)x$ and $\mu(x)$ the Rayleigh quotient. Analyzes the method using the Lanczos polynomials and the symmetric tridiagonal matrix similar to A.

9. /EL/ Hestenes, Magnus R. and William Karush (1951b) "Solutions of $Ax=\lambda Bx$," *J. Res. Nat. Bur. Standards* 49, pp. 471-478.

Extends the Hestenes-Karush (1950) algorithm to the generalized eigenvalue problem.

10. /C/ Hestenes, Magnus R. and Marvin L. Stein (1951) *The Solution of Linear Equations by Minimization,* NAML Report 52-45, December 12, 1951, National Bureau of Standards, Los Angeles, California.

(Reprinted in *J. of Optimization Theory and Applications* 11 (1973), pp. 335-359.) Proposes solving $Ax=b$ by minimizing $(b-Ax)^* H (b-Ax)$, where H is Hermitian positive definite. Studies iterations of the form $x_{k+1}=x_k+\alpha_i p_i$, and derives conditions on α_i and p_i to guarantee convergence. Notes that steepest descent, steepest descent with non-optimal α_i, Gauss-Seidel, SOR, block SOR, n-step methods such as those "investigated by Lanczos, Hestenes, and Stiefel", and other algorithms are special cases.

11. /EL/ Karush, W. (1951) "An Iterative Method for Finding Characteristic Vectors of a Symmetric Matrix," *Pacific J. Math.* 1, pp. 233-248.

    Suggests an algorithm equivalent to taking s steps of the Lanczos algorithm, finding the minimizing eigenvector approximation, and iterating, making reorthogonalization less critical than in the Lanczos algorithm. References Lanczos (1950), but does not really draw the relationship.

12. /ELP/ Rosser, J. B., C. Lanczos, M. R. Hestenes, and W. Karush (1951) "Separation of Close Eigenvalues of a Real Symmetric Matrix," *J. Res. Nat. Bur. Standards* 47, pp. 291-297.

    Solves a difficult 8×8 eigenvalue problem by the Lanczos (1950) algorithm by a "hand computer" in 100 hours (This method "seems best adapted for use by a hand computer using a desk computing machine.") and by the Hestenes and Karush (1951) method (fixed α) on an IBM Card-Programmed Electronic Calculator. ("Considerable time was spent by Karush in becoming familiar with the machine, so that it is difficult to say just how long the computation would require of an experienced operator. Probably 3 or 4 days would be ample.") Suggests polynomial preconditioning to increase the separation of the eigenvalues.

1952

13. /CP/ Hestenes, Magnus R. and Eduard Stiefel (1952) "Methods of Conjugate Gradients for Solving Linear Systems," *J. Res. Nat. Bur. Standards* 49, pp. 409-436.

    "The method of conjugate gradients was developed independently by E. Stiefel of the Institute of Applied Mathematics at Zurich and by M. R. Hestenes with the cooperation of J. B. Rosser, G. Forsythe, and L. Paige of the Institute for Numerical Analysis, National Bureau of Standards. The present account was prepared jointly by M. R. Hestenes and E. Stiefel during the latter's stay at the National Bureau of Standards. The first papers on this method were given by E. Stiefel [1952] and M. R. Hestenes [1951]. Reports on this method were given by E. Stiefel and J. B. Rosser at a Symposium on August 23-25, 1951. Recently, C. Lanczos [1952] developed a closely related routine based on his earlier paper on eigenvalue problem. Examples and numerical tests of the method have been by R. Hayes, U. Hochstrasser, and M. Stein."

    For A symmetric and positive definite: develops conjugate gradients as an iterative method noting that x_{n+1} is often considerably better than x_n although earlier convergence may occur. Gives the $x-r-p$ version of the algorithm and notes that $||x-x^*||$ and $||x-x^*||_{A^{1/2}}$ are monotonically decreasing although the residual norm may oscillate. Gives formulas for obtaining characteristic roots from the recurrences. Proves algebraic and geometric properties of conjugate direction and conjugate gradient algorithms and references Fox, Huskey, and Wilkinson (1948). Gives an algorithm in which the residual norm is monotonic which modifies the x iterates from conjugate gradients. Gives some round-off analysis and recommends smoothing the initial residual. Gives an end correction procedure in case orthogonality is lost. Investigates other normalizations for the direction vectors.

For A symmetric semidefinite: notes that conjugate gradients can obtain a least squares solution.

For general A: uses A^*A-type algorithm.

Also presents the conjugate direction and conjugate gradient algorithms applied to $MAx=Mb$ and gives the examples $M=I$ and $M=A^*$. Shows that conjugate directions with unit vectors applied to a symmetric matrix is equivalent to Gauss elimination. Gives a conjugate direction example in which $\|x-x^*\|$ is monotonically increasing at intermediate steps. Describes a duality between orthogonal polynomials and n-dimensional geometry. Gives the 3-term recurrence relations for the residual polynomials. Notes the relation to the Lanczos (1950) algorithm for computing characteristic polynomials and that the conjugate gradient parameters can be computed by continued fraction expansion of a ratio of polynomials in A. Recommends computational formulas $\alpha=r^Tp/p^TAp$ and $\beta=-r^TAp/p^TAp$. Gives numerical examples and notes that the largest system yet solved involved 90 iterations on 106 difference equations.

14. /CL/ Lanczos, Cornelius (1952) "Solution of Systems of Linear Equations by Minimized Iterations," *J. Res. Nat. Bur. Standards* 49, pp. 33-53.

"The latest publication of Hestenes [1951] and of Stiefel [1952] is closely related to the p,q algorithm of the present paper, although developed independently and from different considerations." "The present investigation is based on years of research concerning the behavior of linear systems, starting with the author's consulting work for the Physical Research Unit of the Boeing Airplane Company, and continued under the sponsorship of the National Bureau of Standards." Applies the Lanczos (1950) algorithm to solving nonsymmetric systems of linear equations by generating a double set of vectors (equivalently, polynomials) $p_k=p_k(A)b$ with leading coefficient 1 so that $\|p_k\|$ is minimal, and $q_k=q_k(A)b$ with constant coefficient 1 so that $\|q_k\|$ is minimal. Shows that the p and p^* sequences are biorthogonal and that the q sequences (saving the vectors) can be used to construct minimal residual solutions for the original system and others involving the same matrix. Advocates complete reorthogonalization or periodic restart to reduce the accumulation of error. Recommends scaling symmetric matrices to make diagonal elements 1 and nonsymmetric matrices to make column norms 1. If A has real nonnegative eigenvalues, recommends a two-phase algorithm, "purifying" the right hand side by Chebyshev polynomial iteration designed to damp out the components corresponding to large eigenvalues and then running the minimized iteration algorithm on the remainder. Notes that this can also be used to give smooth approximate solutions to nearly singular problems by terminating the second phase when a correction vector becomes too large.

15. /CLP/ Stein, Marvin L. (1952) "Gradient Methods in the Solution of Systems of Linear Equations," *J. Res. Nat. Bur. Standards* 48, pp. 407-413.

Reports on numerical experiments on a preconditioned form of steepest descent (but preconditioning is not used for acceleration). Also converts linear system to an eigenvalue problem and applies algorithm of Hestenes and Karush (1951).

16. /C/ Stiefel, Eduard (1952) "Über einige Methoden der Relaxationsrechnung," *Zeitschrift für angewandte Mathematik und Physik* 3, pp. 1-33.

(A version of this paper was presented at the NBS conference in August, 1951.) Surveys Jacobi- and Gauss-Seidel-type methods and steepest descent. Defines "simultaneous relaxation" as adding a linear combination of several vectors to the current guess. Notes that the parameters are easy to calculate if the directions are conjugate. Defines a "n-step iteration" (conjugate gradients) and notes that it can also be used to solve other linear systems with the directions already generated, to invert matrices, and to solve eigenvalue problems as Lanczos (1950) does. Uses the 5-point operator as a model problem, and provides numerical experiments. "Note added in proof: After writing up the present work, I discovered on a visit to the Institute for Numerical Analysis (University of California) that these results were also developed somewhat later by a group there. An internal preliminary report for the National Bureau of Standards was given by M. R. Hestenes in August, 1951 (N.A.M.L. Report 52-9)."

17. /C/ Stiefel, Eduard (1952-53) "Ausgleichung ohne Aufstellung der Gaussschen Normalgleichungen," *Wissenschaftliche Zeitschrift der Technischen Hochschule Dresden* 2, pp. 441-442.

Proposes a conjugate direction algorithm for solving least squares problems which uses $A^T r$ as the initial direction, and keeps the directions AA^T-conjugate. (This algorithm later became known as the LSCG algorithm.)

1953

18. /P/ Forsythe, George E. (1953) "Solving Linear Algebraic Equations Can Be Interesting," *Bull. Amer. Math. Soc.* 59, pp. 299-329.

"With the concept of 'ill-conditioned' systems $Ax=b$ goes the idea of 'preconditioning' them. Gauss and Jacobi made early contributions to this subject [referring to the trick of adding an extra equation to a least squares system]. . . . A convenient means of preconditioning is to premultiply the system with a matrix B, so that one has to solve $BAx=Bb$" (p. 318). Gives two examples: $B=A^T$, giving the normal equations, and B being the operator generated by Gaussian elimination, so that BA is upper triangular.

19. /ELP/ Hestenes, Magnus R. (1953) "Determination of Eigenvalues and Eigenvectors of Matrices," in *Simultaneous Linear Equations and the Determination of Eigenvalues*, ed. L. J. Paige and Olga Taussky, Applied Mathematics Series 29, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., pp. 89-94.

Surveys methods used at NBS for symmetric eigenvalue computation: power algorithm, steepest descent on the Rayleigh quotient, two forms of a "generalization" of the Lanczos (1950) algorithm which uses a preconditioning matrix that commutes with A to obtain the recursion for the characteristic polynomial, and a block method. Also discusses the generalized eigenvalue problem.

20. /C/ Rosser, J. Barkley (1953) "Rapidly Converging Iterative Methods for Solving Linear Equations," in *Simultaneous Linear Equations and the Determination of Eigenvalues*, ed. L. J. Paige and Olga Taussky, Applied Mathematics Series 29, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., pp. 59-64.

Describes the conjugate direction algorithm in its general form, but states that the identity preconditioner is convenient. "Through the use of colloquia and discussion groups, nearly all scientific members of the Institute have made some sort of contribution to the problem. Accordingly, it is impossible to assign complete credit for the results disclosed herein to a single person or a few persons. However, certain members of the staff have given concentrated attention to the problem over an extended period and are primarily responsible for the results noted herein. In alphabetical order, these are G. E. Forsythe, M. R. Hestenes, C. Lanczos, T. Motzkin, L. J. Paige, and J. B. Rosser."

21. /EL/ Rutishauser, H. (1953) "Beiträge zur Kenntnis des Biorthogonalisierungs-Algorithmus von Lanczos," *Zeitschrift für angewandte Mathematik und Physik* 4, pp. 35-56.

Proves that there is a starting vector for the Lanczos algorithm which generates m vectors when the degree of the characteristic polynomial is m. Advocates a method of making the co-diagonal elements small when the eigenvalues are real, thus improving the convergence of algorithms to find eigenvalues of the bidiagonal matrix. Gives bounds for the eigenvalues. Relates the algorithm to a system of differential equations.

22. /P/ Shortley, George (1953) "Use of Tschebyscheff-Polynomial Operators in the Numerical Solution of Boundary-Value Problems," *J. of Appl. Phys.* 24, pp. 392-396.

Uses Chebyshev acceleration of the Jacobi algorithm for solving difference approximations to elliptic partial differential equations.

23. /CEL/ Stiefel, Eduard (1953) "Some Special Methods of Relaxation Technique," in *Simultaneous Linear Equations and the Determination of Eigenvalues*, ed. L. J. Paige and Olga Taussky, Applied Mathematics Series 29, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., pp. 43-48.

Presents the conjugate gradient algorithm as a minimization procedure and gives results of numerical experiments on Laplace's equation on a 3×3 grid and calculation of an Airy stress function for the profile of a dam (139 unknowns, 100 hours of computation on the Zurich Relay Calculator). Notes that the residuals are orthogonal and may be used to calculate eigenvalues. "The resulting procedure is similar to that suggested by Lanczos[1950]."

1954

24. /CP/ Curtiss, J. H. (1954) "A Generalization of the Method of Conjugate Gradients for Solving Systems of Linear Algebraic Equations," *Math. Tables and Aids to Comp.* 8, pp. 189-193.

Develops conjugate gradient algorithm for solving nonsymmetric systems by applying it to $BATA^TB^T$. Explains that $B=I, T=A^{-1}$ gives the Hestenes and Stiefel (1952) algorithm, $B=A^T, T=(A^TA)^{-1}$ gives the Hestenes and Stiefel least squares-type algorithm, and $B=I, T=I$ gives the Craig (1955) algorithm.

25. /C/ Forsythe, A. I. and G. E. Forsythe (1954) "Punched-Card Experiments with Accelerated Gradient Methods for Linear Equations," in *Contributions to the Solution of Systems of Linear Equations and the Determination of Eigenvalues*, ed. Olga Taussky, Applied Mathematics Series 39, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., pp. 55-69.

    Runs the Motzkin-Forsythe algorithm (steepest descent with an occasional accelerating step) on 6×6 examples, concluding that it is twice as fast as consistently underrelaxing steepest descent and much faster than steepest descent alone. Notes that the Hestenes and Stiefel (1952) methods seem to supercede these.

26. /C/ Hayes, R. M. (1954) "Iterative Methods of Solving Linear Problems on Hilbert Space," in *Contributions to the Solution of Systems of Linear Equations and the Determination of Eigenvalues*, ed. Olga Taussky, Applied Mathematics Series 39, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., pp. 71-103.

    Extends the conjugate direction algorithms to Hilbert space and proves linear convergence for the conjugate gradient algorithm for general operators and superlinear convergence for operators of the form $I+K$ where K is completely continuous.

27. /EL/ Stecin, I. M. (1954) "The Computation of Eigenvalues Using Continued Fractions," *Uspekhi matem. nauk* 9 No. 2(60), pp. 191-198.

    Discusses Lyusternik's idea for finding the eigenvalues of a symmetric matrix or operator by transforming a series $c_0/z + c_1/z^2 + \cdots c_n/z^n$ (where $c_k=(A^kb,b)$) to a continued fraction whose coefficients are found by Chebyshev's method of moments. Discusses one particular algorithm for doing this, and notes that this is similar to the method of Lanczos orthogonalization.

28. /SP/ Young, David (1954) "On Richardson's Method for Solving Linear Systems with Positive Definite Matrices," *J. of Math. and Physics* 32, pp. 243-255.

    Gives convergence results and optimal choice of acceleration parameters for Richardson's method. Compares with SOR and gradient methods.

1955

29. /C/ Craig, Edward J. (1955) "The *N*-Step Iteration Procedures," *J. of Math. and Physics* 34, pp. 64-73.

    Discusses a set of conjugate direction methods, including the conjugate gradient algorithm, the algorithm built on A^*A-conjugate directions, and the (new) algorithm built on directions which are A^* times a set of orthogonal vectors. Notes that the last two algorithms can be used on symmetric or nonsymmetric matrices.

30. /P/ Forsythe, G. E. and E. G. Straus (1955) "On Best Conditioned Matrices," *Proceedings of the Amer. Math. Soc.* 6, pp. 340-345.

    Studies the problem of minimizing the 2-norm condition number of T^*AT where A is Hermitian positive definite and T is in a class of regular linear transformations. As a special case, determines the optimal diagonal preconditioning matrix to be the one which makes the resulting diagonal elements equal to 1.

31. /CEP/ Hestenes, Magnus R. (1955) "Iterative Computational Methods," *Communications on Pure and Applied Mathematics* 8, pp. 85-96.

Gives a description of the conjugate gradient algorithm in general form and notes that it can be used to solve singular consistent problems. Discusses the eigenvalue problem, but not Lanczos' algorithm. "The terminology 'conjugate gradient' was suggested by the fact that p_i is the gradient of F, apart from a scale factor, on the linear manifold conjugate to $p_0, p_1, \cdots, p_{i-1}$, that is, orthogonal to $[Ap_0, \cdots, Ap_{i-1}]$."

32. /S/ Sheldon, John W. (1955) "On the Numerical Solution of Elliptic Difference Equations," *Math. Tables and Aids to Comp.* 9, pp. 101-112.

Presents the SSOR algorithm

33. /CL/ Stiefel, E. (1955) "Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme," *Comm. math. helv.* 29, pp. 157-179.

Surveys a family of algorithms for solving linear systems. Views steepest descent as Euler's method on the descent trajectory. Establishes one-to-one correspondence between the family of algorithms and sequences of polynomials satisfying $R_i(0)=1$. Gives the Lanczos polynomials and conjugate gradients as example of such a sequence, with polynomials orthogonal with respect to a discrete distribution function. Studies iterations based on the distribution $\lambda^{\alpha}(1-\lambda)^{\beta}$. Gives numerical results for Poisson equation with constant right-hand side.

1956

34. /EL/ Brooker, R. A. and F. H. Sumner (1956) "The Method of Lanczos for Calculating the Characteristic Roots and Vectors of a Real Symmetric Matrix," *Proc. Inst. Elect. Engrs.* B.103 Suppl., pp. 114-119.

Gives expository treatment of Lanczos algorithm. Recommends Jacobi method for small problems, Lanczos with reorthogonalization for large ones.

35. /CEL/ Crandall, Stephen H. (1956) *Engineering Analysis: A Survey of Numerical Procedures,* McGraw-Hill, New York.

Gives textbook description of conjugate gradient and Lanczos algorithms. "The usefulness of these methods for actual calculation is still being evaluated. . . . There is, however, no denying the mathematical elegance of the methods."

36. /ACP/ Fischbach, Joseph W. (1956) "Some Applications of Gradient Methods," in *Proceedings of the Sixth Symposium in Applied Mathematics (1953),* McGraw-Hill, New York, pp. 59-72.

Discusses and experiments with conjugate gradients for computing the inverse of a matrix, for solving a two-point boundary value problem, and for solving a mildly nonlinear differential equation after a close approximation to the solution is obtained. "All those who have carried out computations by the method of conjugate gradients have observed that the $(N+1)$st step is usually better than the Nth and represents an improvement which overcomes rounding-off error. Frequently $2N$ steps are better than N. . . . One possible way of reducing the error growth is to change the metric (change definition of scalar product) so that the matrix is better conditioned."

37. /CP/ Hestenes, Magnus R. (1956a) "The Conjugate-Gradient Method for Solving Linear Systems," in *Proceedings of the Sixth Symposium in Applied Mathematics (1953)*, McGraw-Hill, New York, pp. 83-102.

Derives conjugate direction and conjugate gradient algorithms in general form, minimizing a function with an arbitrary inner product matrix, and having a preconditioning matrix. Notes that the conjugate gradient parameters can be bounded in terms of generalized eigenvalues. Discusses the standard conjugate gradient algorithm and the minimum error norm form. Shows that every n-step iterative method can be reproduced by a conjugate direction method. "From a mathematical point of view [the original Hestenes and Stiefel algorithm] represents the general case in the sense that every conjugate gradient algorithm can be reduced to this form by a change of variable or by a simple change of the original system to be solved." Notes that no essential changes are required to extend to Hilbert space.

38. /C/ Hestenes, Magnus R. (1956b) "Hilbert Space Methods in Variational Theory and Numerical Analysis," in *Proceedings of the International Congress of Mathematicians 1954* 3, North-Holland, Amsterdam, pp. 229-236.

Studies properties of quadratic forms in Hilbert space. Describes conjugate gradients as a minimization method on the error function, summarizing results of Hayes (1954).

39. /CEL/ Lanczos, Cornelius (1956) *Applied Analysis,* Prentice-Hall, Englewood Cliffs, New Jersey.

Discusses use of the p,q Lanczos (1950) algorithm for finding eigenvalues and eigenvectors. Notes that the large eigenvalues are approximated quickly, and the small eigenvalues could be determined by "preliminary inversion of the matrix." Suggests use of Chebyshev polynomial transformation of the matrix to determine eigenvalues in an intermediate range.

1957

40. /C/ Stiefel, E. (1957) "Recent Developments in Relaxation Techniques," in *Proceedings of the International Congress of Mathematicians 1954* 1, North-Holland, Amsterdam, pp. 384-391.

Defines a "relaxation process" as one which reduces a measure of the error at each step. Notes that for symmetric positive definite matrices, Gauss-Seidel, Gauss elimination (considered as an iteration), and gradient methods are relaxation processes. Develops the optimal polynomial property for conjugate gradients.

1958

41. /AEL/ Gregory, R. T. (1958) "Results Using Lanczos' Method for Finding Eigenvalues of Arbitrary Matrices," *J. Soc. Industr. Appl. Math* 6, pp. 182-188.

Uses Lanczos (1950) algorithm for complex non-Hermitian matrices with double precision arithmetic, scaling of vectors, and full re-orthogonalization.

42. /CLP/ Lanczos, C. (1958) "Iterative Solution of Large-Scale Linear Systems," *J. Soc. Industr. Appl. Math* 6, pp. 91-109.

Discusses the effect of ill-conditioning and right-hand-side measurement errors on the accuracy of solutions to linear systems with symmetric coefficient matrices. Analyzes nonsymmetric ones through the symmetric system of size $2n$. Estimates largest eigenvalue by refinement of power method, and scales the matrix by it. Then applies iteration based on Chebyshev polynomials and matrix of dimension $n+2$.

43. /CEL/ Stiefel, Eduard L. (1958) "Kernel Polynomials in Linear Algebra and Their Numerical Applications," in *Further Contributions to the Solution of Simultaneous Linear Equations and the Determination of Eigenvalues,* Applied Mathematics Series 49, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., pp. 1-22.

Derives relaxation algorithms by considering various sets of polynomials with value 1 at zero. Recommends a two stage process for ill-conditioned systems: filter out error components corresponding to a large but clustered set of eigenvalues and then apply conjugate gradients to remove components corresponding to a few small eigenvalues. As an example, solves Laplace's equation with constant right-hand side on a 10×10 grid with 11 Chebyshev iterations on [2,8] and 2 conjugate gradient steps, getting 4 orders of magnitude reduction in the error. Recommends solving nonlinear systems by the change of variables $AA^*y=b$. Applies kernel polynomials to the problem of eigenvalue estimation, obtaining the Lanczos (1950) algorithm, among others. "As it stands, Lanczos' algorithm can only be successful for low-order matrices with nicely separated eigenvalues. For larger matrices the rounding-off errors destroy quickly the orthogonality of the vectors. As in solving linear equations, it is necessary to find for such matrices a suitable *combination of the methods* available." Discusses polynomial transformations to emphasize certain ranges of the spectrum, and applies the Lanczos algorithm to the transformed matrix. Discusses the generation of orthogonal polynomials by the quotient-difference algorithm, including the variant corresponding to the Lanczos algorithm.

44. /AEL/ Wilkinson, J. H. (1958) "The Calculation of Eigenvectors by the Method of Lanczos," *Computer J.* 1, pp. 148-152.

Uses reorthogonalization on symmetric Lanczos and reorthogonalization plus double precision on unsymmetric Lanczos. Notes that the latter is a very powerful algorithm.

1959

45. /C/ Altman, Mieczyslaw (1959) "On the Convergence of the Conjugate Gradient Method for Non-Bounded Linear Operators in Hilbert Space," in *Approximation Methods in Functional Analysis,* Lecture Notes, California Institute of Technology, pp. 33-36.

Proves convergence of conjugate gradients for a self-adjoint, positive definite linear operator, with domain a dense linear space, satisfying $(Au,u)\geq k(u,u)$ for some positive constant k and all u in the domain.

46. /N/ Davidon, W. C. (1959) *Variable Metric Method for Minimization,* Report ANL-5990, Argonne National Laboratory, Argonne, Illinois.

Derives a method (the DFP method, further developed in Fletcher and Powell (1963)) meant to "improve the speed and accuracy with which the minima of functions can be evaluated numerically" compared to such methods as conjugate gradients, steepest descent, and Newton-Raphson. Proposes guessing the inverse Hessian matrix (symmetric and positive definite), and generating a search direction equal to this matrix times the negative gradient. Uses either this direction plus an orthogonal correction, or a line search along this direction, to determine the next iterate. Then modifies the inverse Hessian approximation according to the Quasi-Newton condition using a rank-one or rank-two update. Discusses the initial Hessian approximation and the incorporation of linear constraints. Discusses in an appendix a simplified rank-one updating procedure.

47. /ACP/ Engeli, M., Th. Ginsburg, H. Rutishauser, and E. Stiefel (1959) *Refined Iterative Methods for Computation of the Solution and the Eigenvalues of Self-Adjoint Boundary Value Problems,* Birkhauser Verlag, Basel/Stuttgart.

Stiefel: Solves self-adjoint partial differential equations by variational formulation, not by differential equation itself.

Rutishauser: Surveys gradient methods (Richardson one and two, steepest descent, Frankel (second order Richardson), Chebychev, hypergeometric relaxation, conjugate gradients, conjugate residual). Considers "combined methods," including:

1. conjugate gradients with Chebyshev (attributed to Lanczos (1952)): smooth the residual with Chebyshev polynomial over range of high eigenvalues, then apply conjugate gradients. It is noted that the high eigenvalues are "reactivated" by conjugate gradients, and the method is not recommended.

2. "Replace the system given by another system with the same solution but with a coefficient matrix A of smaller P-condition number." (P-condition number = condition number in the 2-norm.) Polynomial conditioning is explicitly considered and is attributed to Stiefel (1958) in the case of eigencomputation. The iterations are called "inner" for the polynomial and "outer" for conjugate gradients, and Chebyshev-conjugate gradients is recommended.

Notes that methods such as those of Richardson and Frankel can also be used for eigencomputations. Gives the conjugate gradient tridiagonal matrix in nonsymmetric 3-term form. Notes round-off difficulties, and recommends proceeding more or less than n steps, as long as the residual remains small. Recommends comparing the approximations from two tridiagonal matrices (same initial vector, different number of steps, or different initial vector) to determine convergence. Also discusses determining the eigenvalues of the original matrix from the conjugate gradient-Chebyshev method.

Ginsburg: Gives results of numerical experiments. For a finite difference problem with 70 variables, needs approximately 10 conjugate gradient iterations with preconditioning by a 10th-order Chebyshev polynomial. Compares with steepest descent, conjugate gradient, and other methods on this and other examples.

Engeli: Surveys relaxation methods.

Conclusions: For moderate condition problems, use relaxation. For bad conditioning, use conjugate gradients or conjugate gradients-Chebyshev with recursive residuals. Recommends conjugate gradients over conjugate gradients-Chebyshev unless some low eigenvalues are needed.

48. /AC/ Läuchli, Peter (1959) "Iterative Lösung und Fehlerabschätzung in der Ausgleichsrechnung," *Zeitschrift für angewandte Mathematik und Physik* 10, pp. 245-280.

Develops conjugate gradients and other relaxation methods for overdetermined linear systems. Notes that finding the point in an n-dimensional subspace of R^m (spanned by the columns of C) which is closest to a point l is equivalent to solving $C^T Cx = C^T l$, but that the problem can also be formulated in terms of a basis B for the null space of C^T, representing x implicitly as $B^Tx+b=0$ (for some vector b) and avoiding normal equations. Uses Synge's method of the hypercircle to find upper and lower bounds on the sum of squared residuals. Notes that the inverse matrix can be constructed by an update at each conjugate gradient step. Provides numerical examples.

1960

49. /C/ Beckman, F. S. (1960) "The Solution of Linear Equations by the Conjugate Gradient Method," in *Mathematical Methods for Digital Computers*, ed. Anthony Ralston and Herbert S. Wilf, Wiley, New York.

    Derives conjugate gradients as a conjugate direction method including flow chart, comments on error analysis, etc.

50. /CP/ Frank, Werner L. (1960) "Solution of Linear Systems by Richardson's Method," *J. Assoc. Comput. Mach.* 7, pp. 274-286.

    Follows Stiefel (1958) in using Chebyshev acceleration on a partial interval; then applies conjugate gradients. Tests the algorithm on a 50×50 matrix tri(-1,2,-1) with (1,1) element modified to 1. Needs 20 conjugate gradient iterations (instead of the theoretical termination in 5) to get 5 digits of accuracy; requires the full 50 if conjugate gradients is used alone. Required 46 conjugate gradient iterations to solve 5 point difference equations for $n=361$.

51. /P/ Householder, A. S. and F. L. Bauer (1960) "On Certain Iterative Methods for Solving Linear Systems," *Numer. Math.* 2, pp. 55-59.

    Discusses "methods of projection" which iterate $x=x+Yp$, where the columns of Y span a subspace and p is chosen so that the error decreases. Notes that steepest descent and relaxation techniques both fit into this framework, but does not mention conjugate gradients.

52. /AC/ Livesley, R. K. (1960) "The Analysis of Large Structural Systems," *Computer J.* 3, pp. 34-39.

    Tries to apply conjugate gradients to an ill-conditioned system in structural analysis. Finds conjugate gradients ineffective because it requires n matrix multiplications and thus n tape scans, and "rounding errors show a tendency to build up to such an extent that the solution after N steps is often a worse approximation to the correct solution than the starting point." "The method was therefore abandoned in favour of an elimination process."

53. /EP/ Osborne, E. E. (1960) "On Pre-Conditioning of Matrices," *J. Assoc. Comput. Mach.* 7, pp. 338-345.

    Constructs a sequence of diagonal similarity transformations to scale an irreducible matrix to increase the smallest eigenvalue relative to the matrix norm so that the eigensystem can be more easily determined.

54. /PS/ Varga, Richard S. (1960) "Factorization and Normalized Iterative Methods," in *Boundary Problems in Differential Equations*, ed. Rudolph E. Langer, University of Wisconsin Press, Madison, pp. 121-142.

    "The main purpose of this article is to introduce a class of iterative methods which depend upon the direct solution of matrix equations involving matrices more general than tridiagonal matrices." Assumes that A is Stieltjes. Introduces the idea of regular splitting. Given a regular splitting, accelerates by overrelaxation, Chebyshev semi-iteration, Peaceman-Rachford (1955) algorithm, or Douglas-Rachford (1956) algorithm. Suggests normalizing factors to have unit diagonal for computational efficiency. Discusses effectiveness of successive line overrelaxation. Proposes approximate factorization of A, keeping the factors as sparse as A (the algorithm that has come to be known as incomplete Cholesky factorization with no extra diagonals). Shows that this yields a regular splitting for the 5-point operator.

55. /PS/ Wachspress, E. L. and G. J. Habetler (1960) ''An Alternating-Direction-Implicit Iteration Technique,'' *J. Soc. Industr. Appl. Math.* 8, pp. 403-424.

    ''Conditions'' a matrix by diagonal scaling before applying ADI.

56. /CN/ Zoutendijk, G. (1960) *Methods of Feasible Directions,* Elsevier, Amsterdam.

    Uses conjugate directions to construct a finitely terminating quadratic programming algorithm.

1961

57. /AEL/ Causey, R. L. and R. T. Gregory (1961) ''On Lanczos' Algorithm for Tridiagonalizing Matrices,'' *SIAM Rev.* 3, pp. 322-328.

    Discusses biorthogonal reduction to tridiagonal form. Distinguishes between fatal and nonfatal instances when the inner product between the left and right vectors vanish.

58. /S/ D'Yakonov, E. G. (1961) ''An Iteration Method for Solving Systems of Finite Difference Equations,'' *Soviet Math. Dokl.* 2, pp. 647-650.

    (Dokl. Akad. Nauk SSSR 138, pp. 271-274.) Analyzes the iteration $Mx_{k+1}=Mx_k-\omega(Ax_k-b)$ where A is a finite difference approximation to an elliptic operator over the unit square and M represents several iterations of the ADI operator for the Laplacian. Gives a work estimate of $O(n \ln^2 n^{-1/2}) \ln \varepsilon$ to solve the problem with precision ε.

59. /PS/ Golub, Gene H. and Richard S. Varga (1961) ''Chebyshev Semi-Iterative Methods, Successive Overrelaxation Iterative Methods, and Second Order Richardson Iterative Methods, Parts I and II,'' *Numer. Math.* 3, pp. 147-156, 157-168.

    Compares the rates of convergence of the three iterative methods of the title. Proposes applying Chebyshev acceleration to two-cyclic matrices resulting from block preconditionings, such as those derived from the block SOR splitting of Arms, Gates, and Zondek (*SIAM J.* 4, 1956, pp. 220-229). Gives applications to partial difference equations.

60. /PS/ Habetler, G. J. and E. L. Wachspress (1961) ''Symmetric Successive Overrelaxation in Solving Diffusion Difference Equations,'' *Math. of Comp.* 15, pp. 356-362.

    Uses Chebyshev acceleration on Sheldon's SSOR algorithm (*J. Assoc. Comput. Mach.* 6, 1959, pp. 494-505). Shows SSOR not effective in diffusion calculations in nuclear reactor theory if the grids are too irregular. Gives algorithm to estimate parameters.

61. /C/ Martin, D. W. and G. J. Tee (1961) ''Iterative Methods for Linear Equations with Symmetric Positive Definite Matrix,'' *Computer J.* 4, pp. 242-254.

    Surveys stationary iterative methods, steepest descent, and conjugate gradients including previous numerical results. Concludes that ''no single method is to be recommended for universal applications.''

62. /S/ Oliphant, Thomas A. (1961) "An Implicit, Numerical Method for Solving Two-Dimensional Time-Dependent Diffusion Problems," *Quarterly of Appl. Math.* 19, pp. 221-229.

Proposes an iterative method for nine-point finite difference approximations, using a partial factorization of the difference matrix as a splitting. Applies the algorithm to linear and nonlinear problems.

63. /EL/ Rollett, J. S. and J. H. Wilkinson (1961) "An Efficient Scheme for the Co-diagonalization of a Symmetric Matrix by Givens' Method in a Computer with a Two-level Store," *Computer J.* 4, pp. 177-180.

Notes that the resulting bidiagonal matrix for their algorithm is the same as that from the Lanczos (1950) algorithm.

64. /EL/ Strachey, C. and J. G. F. Francis (1961) "The Reduction of a Matrix to Codiagonal Form by Eliminations," *Computer J.* 4, pp. 168-176.

Notes that the Lanczos (1950) method is equivalent to an elimination method for reduction of a Hessenberg matrix to tridiagonal form.

1962

65. /C/ Antosiewicz, Henry A. and Werner C. Rheinboldt (1962) "Numerical Analysis and Functional Analysis," in *Survey of Numerical Analysis*, ed. John Todd, McGraw-Hill, New York, pp. 485-517 (Ch. 14).

Presents conjugate directions for linear self-adjoint positive definite operators on Hilbert space and proves a convergence rate.

66. /AC/ Bothner-By, Aksel A. and C. Naar-Colin (1962) "The Proton Magnetic Resonance Spectra of 2,3-Disubstituted n-Butanes," *J. of the ACS* 84, pp. 743-747.

Analyzes chemical spectra by solving a least squares problem with conjugate gradients.

67. /ACN/ Feder, Donald P. (1962) "Automatic Lens Design with a High-Speed Computer," *J. of the Optical Soc. of Amer.* 52, pp. 177-183.

Suggests conjugate gradients or DFP methods, among others, to minimize a merit function in lens design.

68. /S/ Oliphant, Thomas A. (1962) "An Extrapolation Procedure for Solving Linear Systems," *Quarterly of Appl. Math.* 20, pp. 257-265.

Generalizes the method of Oliphant (1961) to five-point operators, and allows partial factorizations of a modified difference matrix.

69. /C/ Petryshyn, W. V. (1962) "Direct and Iterative Methods for the Solution of Linear Operator Equations in Hilbert Space," *Trans. AMS* 105, pp. 136-175..

Derives minimum error method and, from it, other algorithms. Does not use the extra matrices for preconditioning.

70. /N/ Powell, M. J. D. (1962) "An Iterative Method for Finding Stationary Values of a Function of Several Variables," *Computer J.* 5, pp. 147-151.

Proposes a method which, given a starting point x_0, finds a minimizer in one direction, x_1, then minimizes in the $n-1$ dimensional hyperplane through x_1 orthogonal to the first direction, giving x_2. Then the minimizer is on the line between x_0 and x_2. The method is "not unlike the conjugate gradient method of Hestenes and Stiefel (1952)."

71. /EL/ Wilkinson, J. H. (1962) "Instability of the Elimination Method of Reducing a Matrix to Tri-Diagonal Form," *Computer J.* 5, pp. 61-70.

Relates the Lanczos (1950) algorithm to Hessenberg's method (1941 Ph.D. thesis) applied to a lower Hessenberg matrix, reducing it to tridiagonal form.

1963

72. /CEL/ Faddeev, D. K. and V. N. Faddeeva (1963) *Computational Methods of Linear Algebra,* W. H. Freeman and Co., San Francisco, California.

(Translated by Robert C. Williams from 1960 publication of State Publishing House for Physico-Mathematical Literature, Moscow.) Discusses in Chapter 4 the "method of orthogonalization of successive iterations" for finding eigenvalues of matrices, which, in the symmetric case, is the Lanczos (1950) algorithm. Discusses in Chapter 6 how to continue the algorithm for symmetric and nonsymmetric matrices in case it terminates in fewer than n steps. Discusses the use of the "A-minimal iteration algorithm," the "A-biorthogonal algorithm," steepest descent, s-dimensional steepest descent, and conjugate direction algorithms for solving linear systems.

73. /N/ Fletcher, R. and M. J. D. Powell (1963) "A Rapidly Convergent Descent Method for Minimization," *Computer J.* 6, pp. 163-168.

Derives the Davidon-Fletcher-Powell (DFP) algorithm for minimizing non-quadratic functions and accumulating an approximate Hessian matrix. References Hestenes and Stiefel (1952).

74. /CL/ Fridman, V. M. (1963) "The Method of Minimum Iterations with Minimum Errors for a System of Linear Algebraic Equations with a Symmetric Matrix," *USSR Comp. Math. and Math. Phys.* 2, pp. 362-363.

Derives a conjugate gradient method (from the Lanczos perspective) which minimizes the 2-norm of the error over the subspace $Ar^{(0)}, A^2r^{(0)}, \cdots$.

75. /C/ Ginsburg, Theo (1963) "The Conjugate Gradient Method," *Numer. Math.* 5, pp. 191-200.

(The Handbook Series Linear Algebra conjugate gradient algorithm.) Uses the 3-term recurrence version of the conjugate gradient algorithm.

76. /C/ Khabaza, I. M. (1963) "An Iterative Least-Square Method Suitable for Solving Large Sparse Matrices," *Computer J.* 6, pp. 202-206.

Proposes the s-dimensional steepest descent algorithm applied to minimization of the norm of the residual for solving linear systems. Does not recompute the parameters in subsequent iterations unless the residual begins to increase. Notes superiority to conjugate gradients and SOR on some test problems.

77. /ACPS/ Wachspress, Eugene L. (1963) "Extended Application of Alternating Direction Implicit Iteration Model Problem Theory," *J. Soc. Industr. Appl. Math.* 11, pp. 994-1016.

Uses ADI applied to the model problem as a preconditioner for conjugate gradients applied to more general problems. Gives some discussion of convergence rate as a function of mesh spacing. References Lanczos (1952) rather than Hestenes and Stiefel. References Engeli et al. (1959) for other examples of "compound iteration."

1964

78. /ACP/ Dufour, H. M. (1964) "Resolution des Systemes Lineaires par la Methode des Residus Conjugues," *Bulletin Géodésique* 71, pp. 65-87.

Derives the minimum residual and conjugate gradient algorithms and proposes their use for symmetric positive definite systems, for linear least squares problems, for least squares subject to equality constraints, and for systems resulting from block elimination of a 2×2 block matrix, leading to a Schur complement of the form $C-B^*A^{-1}B$ as the matrix in the problem. Discusses preconditioning when an approximate inverse is available. Applies the method to problems in geodesy.

79. /PS/ Ehrlich, Louis W. (1964) "The Block Symmetric Successive Overrelaxation Method," *J. Soc. Industr. Appl. Math.* 12, pp. 807-826.

Uses Chebyshev acceleration on block SSOR. Estimates rate of convergence and gives numerical results.

80. /N/ Fletcher, R. and C. M. Reeves (1964) "Function Minimization by Conjugate Gradients," *Computer J.* 7, pp. 149-154.

Generalizes conjugate gradients to nonquadratic functions by adding line searches and by taking the current gradient to be the current residual. Quadratic termination is obtained without evaluating or approximating the Hessian matrix.

81. /S/ Gunn, James E. (1964a) "The Numerical Solution of $\nabla . a\nabla u=f$ by a Semi-Explicit Alternating-Direction Iterative Technique," *Numer. Math.* 6, pp. 181-184.

Proposes and analyzes the iteration $Mx_{n+1}=Mx_n-\omega(Ax_n-b)$ where M is one step of the Peaceman-Rachford ADI iteration for the discretization of the desired operator $\nabla . a\nabla$ and the domain is rectangular. Obtains a work estimate of $O\,(h^{-2}\log h^{-1}\log \varepsilon^{-1})$ to reduce the error by ε.

82. /S/ Gunn, James E. (1964b) "The Solution of Elliptic Difference Equations by Semi-Explicit Iterative Techniques," *SIAM J. Numer. Anal.* 2(Series B), pp. 24-45.

Proposes and analyzes the iteration $Mx_{n+1}=Mx_n-\omega(Ax_n-b)$ where M is one step of the Peaceman-Rachford ADI iteration (variable ω) for the discrete Laplacian operator (i.e., not the matrix A), the elliptic operator is not necessarily symmetric, and the domain is rectangular. Uses Chebyshev acceleration and second-order Richardson and obtains an improved convergence result over Gunn (1964). Applies the algorithm to mildly nonlinear problems.

83. /CEL/ Householder, Alston S. (1964) *The Theory of Matrices in Numerical Analysis,* Blaisdell Publishing Co., New York.

"The Lanczos algorithm is well known in the theory of orthogonal polynomials, but Lanczos (1950) seems to have been the first to apply it to the reduction of matrices (p.28, Dover edition)." Develops Lanczos tridiagonalization in matrix form; discusses Lanczos polynomials.

84. /EL/ Parlett, Beresford (1964) "The Development and Use of Methods of LR type," *SIAM Rev.* 6, pp. 275-295.

Notes that Henrici observed that the first diagonal of the QD scheme can be found by the Lanczos (1950) algorithm; thus, QD links the power method to Lanczos' method.

85. /CN/ Powell, M. J. D. (1964) "An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives," *Computer J.* 7, pp. 155-162.

Proposes an algorithm which uses n line searches per iteration to generate a new direction. Shows that, for a quadratic function, the directions are conjugate. Proposes a modification in case the n line search directions become linearly dependent. Gives numerical examples.

86. /C/ Pyle, L. Duane (1964) "Generalized Inverse Computations Using the Gradient Projection Method," *J. Assoc. Comput. Mach.* 11, pp. 422-428.

    Notes that a Gram-Schmidt-based algorithm for computing generalized inverses is a conjugate direction method if the matrix is square.

87. /N/ Shah, B. V., R. J. Buehler, and O. Kempthorne (1964) "Some Algorithms for Minimizing a Function of Several Variables," *J. Soc. Industr. Appl. Math.* 12, pp. 74-92.

    Introduces Partan, a method with quadratic termination in $2n-1$ steps or fewer, which generates conjugate directions. Includes preconditioning matrix in the formulation.

1965

88. /CN/ Broyden, C. G. (1965) "A Class of Methods for Solving Nonlinear Simultaneous Equations," *Math. of Comp.* 19, pp. 577-593.

    Develops a family of algorithms based on satisfying the quasi-Newton condition and using rank-one or rank-two updates to the approximate derivative matrix at each iteration. Proposes three update formulas. Proposes either a step-size of one, or using the norm of the residual in a criterion for termination of the line search. Gives an Algol program and numerical results on ten test problems.

89. /AC/ Campbell, William J. (1965) "The Wind-Driven Circulation of Ice and Water in a Polar Ocean," *J. of Geophysical Research* 70, pp. 3279-3301.

    Solves nonlinear equations with 780 variables using conjugate gradients on a linear system at each iteration.

90. /CN/ Fletcher, R. (1965) "Function Minimization without Evaluating Derivatives - A Review," *Computer J.* 8, pp. 33-41.

    Reviews conjugate direction methods of a 1962 paper of Smith, Powell (1964) and a 1964 paper of Davies, Swann, and Campey.

91. /L/ Golub, G. and W. Kahan (1965) "Calculating the Singular Values and Pseudo-Inverse of a Matrix," *SIAM J. Numer. Anal.* 2(Series B), pp. 205-224.

    Uses Lanczos' observation that the singular values of a matrix are the eigenvalues of a matrix of dimension $n+m$ with zeros on the block diagonal and A and A^* off the diagonal. Generates the bidiagonal form from Householder transformations or from the Lanczos (1950) algorithm.

92. /CN/ Nashed, M. Z. (1965) "On General Iterative Methods for the Solutions of a Class of Nonlinear Operator Equations," *Math. of Comp.* 19, pp. 14-24.

    Gives a class of iterative methods for operators in Hilbert space and shows that conjugate gradients and others are first-order approximations to these methods.

93. /AN/ Paiewonsky, Bernard (1965) "Optimal Control: A Review of Theory and Practice," *AIAA J.* 3, pp. 1985-2006.

    Surveys control problems and nonlinear optimization methods.

94. /CEL/ Vorobyev, Yu V. (1965) *Method of Moments in Applied Mathematics,* Gordan and Breach Science Pub., New York.

(Translated from the Russian by B. Seckler.) Discusses the Lanczos algorithm for generating an orthogonal basis and and a sequence of orthogonal polynomials for symmetric matrices. Notes its use in finding eigenvalues, and discusses Lyusternik's idea of converting a certain infinite series to a continued fraction in order to use moments to determine the coefficients of the same orthogonal polynomials that Lanczos used. References L. A. Lyusternik, "Solution of Linear Algebraic Problems by the Method of Continued Fractions," *Transactions of a Seminar on Functional Analysis,* No. 2, Voronezh (1956), pp. 85-90. Notes that the "point of departure" for Lanczos (1950) and (1952) and for Hestenes and Stiefel (1952) and Lyusternik was "the Chebyshev-Markov classical scalar problem of moments for the quadratic functional (Ax,x)," and that "the methods of Lanczos and Lyusternik were subsequently extended to completely continuous self-adjoint operators" by Karush (1952) and Stecin (1954).

95. /AN/ Wilde, D. J. (1965) "A Review of Optimization Theory," *Indust. and Eng. Chem.* 57 no. 8, pp. 19-31.

Mentions conjugate gradients and other methods.

96. /CEL/ Wilkinson, J. H. (1965) *The Algebraic Eigenvalue Problem,* Clarendon Press, Oxford.

Advocates use of Lanczos (1950) algorithm with double precision arithmetic and complete reorthogonalization. Restarts with different initial vectors if the size of the new vector deteriorates in the nonsymmetric case.

1966

97. /EN/ Bradbury, W. W. and R. Fletcher (1966) "New Iterative Methods for Solution of the Eigenproblem," *Numer. Math.* 9, pp. 259-267.

Uses conjugate gradient and DFP algorithms to solve the generalized symmetric eigenproblem by minimizing the Rayleigh quotient. Notes that the line searches can be performed exactly. Renormalizes at each step to keep the infinity norm of the iterate equal to one. Reports faster convergence with conjugate gradients except on very ill-conditioned problems, but both methods are slower than QR if many eigenvalues are desired.

98. /PS/ D'Yakonov, Ye. G. (1966) "The Construction of Iterative Methods Based on the Use of Spectrally Equivalent Operators," *USSR Comp. Math. and Math. Phys.* 6, No. 1, pp. 14-46.

(Zh. vȳchisl. Mat. mat. Fiz. 6, No. 1, pp. 12-34.) Uses spectrally equivalent operators in a Richardson iterative algorithm and analyzes convergence.

99. /CELP/ Kaniel, Shmuel (1966) "Estimates for Some Computational Techniques in Linear Algebra," *Math. of Comp.* 95, pp. 369-378.

Develops convergence bounds for conjugate gradients in terms of Chebyshev polynomials and the condition number of the matrix. Develops bounds for Lanczos (1950) method eigenvalues in terms of condition number and separations. Notes that results extend to Hilbert space. Results corrected in Belford and Kaufman (1974).

100. /EL/ Lehmann, N. J. (1966) "Zer Verwendung optimaler Eigenwerteingrenzungen bei der Lösung symmetrischer Matrizenaufgaben," *Numer. Math.* 8, pp. 42-55.

Develops a previous idea of using a set of Rayleigh quotients to estimate eigenvalues to the special case where the test vectors are those from the Lanczos (1950) recursion and determines inclusion intervals for the largest. Applies the algorithm to *tri* (−1,2,−1) for n =30. Gets good estimates for 4 eigenvalues after 8 iterations.

101. /N/ Mitter, S., L. S. Lasdon, and A. D. Waren (1966) "The Method of Conjugate Gradients for Optimal Control Problems," *Proc. IEEE* 54, pp. 904-905.

Notes that the Fletcher-Reeves (1964) method also applies in function space.

102. /AN/ Pitha, J. and R. Norman Jones (1966) "A Comparison of Optimization Methods for Fitting Curves to Infrared Band Envelopes," *Canadian J. of Chemistry* 44, pp. 3031-3050.

Concludes that DFP is more effective than a nonlinear conjugate gradient method.

103. /CEP/ Wachspress, Eugene L. (1966) *Iterative Solution of Elliptic Systems and Applications to the Neutron Diffusion Equations of Reactor Physics,* Prentice-Hall, Englewood Cliffs, New Jersey.

In Chapter 5, derives the Lanczos (1950) algorithm and "combined" algorithms (e.g., Lanczos-Chebyshev) in a way similar to Engeli et al. (1959). Notes that the algorithms can be applied to a product of two symmetric matrices. Derives the Chebyshev algorithm for real eigenvalues and for eigenvalues bounded by an ellipse in the complex plane. Discusses Lanczos' eigenvalue algorithm with initial filtering and with a polynomial in A as the operator. In Chapter 6, discusses premultiplication of the linear system by a matrix, and applying the Lanczos or Chebyshev algorithm to the transformed system. Uses ADI preconditioning as an example. Gives a rate of convergence estimate for the model problem ADI preconditioned algorithm. In Chapter 9, derives a multigrid algorithm, relating the idea of contracting the basis to Lanczos projection, and performs numerical experiments indicating improvement over the Golub-Varga two-cyclic version of the Chebyshev algorithm and over SOR.

104. /AN/ Wilson, Robert (1966) "Computation of Optimal Controls," *J. of Math. Anal. and Applics.* 14, pp. 77-82.

Changes a constrained optimization problem to an unconstrained dual problem, decomposes it into subproblems, and applies conjugate gradients.

105. /AN/ Young, P. C. (1966) "Parameter Estimation and the Method of Conjugate Gradients," *Proc. IEEE* 54, pp. 1965-1967.

Uses Mitter, Lasdon, Waren (1966) version of Fletcher-Reeves (1964) algorithm for real-time process parameter estimation. "Unfortunately, the excellent characteristics of the conjugate gradients approach . . . are not maintained as the level of additive noise is increased. Considerable data averaging or 'smoothing' becomes necessary even for low noise levels, and this tends to destroy the real-time nature of the algorithm."

1967

106. /CN/ Broyden, C. G. (1967) "Quasi-Newton Methods and Their Application to Function Minimization," *Math. of Comp.* 21, pp. 368-381.

Further develops the algorithms in Broyden (1965), focusing on rank-two updates, function minimization, and linear systems with symmetric matrices.

107. /CNP/ Daniel, J. W. (1967a) "The Conjugate Gradient Method for Linear and Nonlinear Operator Equations," *SIAM J. Numer. Anal.* 4, pp. 10-26.

Presents convergence rates for the conjugate gradient iteration for bounded linear operators with bounded inverse. Discusses conjugate gradients in full generality, with preconditioning matrix and inner product matrix. Suggests using Laplacian operator to precondition second-order linear elliptic partial differential equations. Nonlinear results corrected in Daniel (1970) and Cohen (1972).

108. /N/ Daniel, James W. (1967b) "Convergence of the Conjugate Gradient Method with Computationally Convenient Modifications," *Numer. Math.* 10, pp. 125-131.

Proves convergence and correct asymptotic rate constant for nonlinear conjugate gradients with inexact line searches. Replaces direction vector parameter by formulas which do not involve the second derivative matrix. Nonlinear results corrected in Daniel (1970) and Cohen (1972).

109. /C/ Forsythe, George E. (1967) "Today's Computational Methods of Linear Algebra," *SIAM Rev.* 9, pp. 489-515.

Gives conjugate gradients credit to Lanczos (1950) and Hestenes and Stiefel (1952). Notes that in 1953, the stability of conjugate gradients was much better understood than that of Gauss elimination.

110. /N/ Lasdon, L. S., S. K. Mitter, and A. D. Waren (1967) "The Conjugate Gradient Method for Optimal Control Problems," *IEEE Trans. on Auto. Control* AC-12, pp. 132-138.

Derives a function space version of the nonlinear conjugate gradient method. Proves convergence if function is bounded below, continuously Frechet differentiable, and the second Frechet derivative is bounded.

111. /AN/ Pitha, J. and R. Norman Jones (1967) "An Evaluation of Mathematical Functions to Fit Infrared Band Envelopes," *Canadian J. of Chemistry* 45, pp. 2347-2352.

Solves nonlinear least squares problems by Levenberg-Marquardt with conjugate gradients.

112. /AN/ Sinnott, Jr., J. F. and D. G. Luenberger (1967) "Solution of Optimal Control Problems by the Method of Conjugate Gradients," in *1967 Joint Automatic Control Conference, Preprints of Papers*, Lewis Winner, New York, pp. 566-574.

Develops the conjugate gradient algorithm for minimization subject to linear equality constraints and applies it to control problems, obtaining superlinear convergence.

113. /CN/ Zangwill, Willard I. (1967) "Minimizing a Function Without Calculating Derivatives," *Computer J.* 10, pp. 293-296.

Modifies the algorithm of Powell (1964) to handle the case where the directions fail to be linearly independent and proves convergence for strictly convex functions. Proposes an alternate method.

1968

114. /PS/ Dupont, Todd (1968) "A Factorization Procedure for the Solution of Elliptic Difference Equations," *SIAM J. Numer. Anal.* 5, pp. 753-782.

Extends the Dupont, Kendall, and Rachford (1968) results to different boundary conditions and mildly nonlinear problems, obtaining a work estimate of $O(h^{-d-1/2}\log \varepsilon^{-1})$ for d-dimensional problems to reduce the error by ε.

115. /PS/ Dupont, Todd, Richard P. Kendall, and H. H. Rachford, Jr. (1968) "An Approximate Factorization Procedure for Solving Self-Adjoint Elliptic Difference Equations," *SIAM J. Numer. Anal.* 5, pp. 559-573.

Analyzes the iteration $Mx_{n+1}=Mx_n-\omega(Ax_n-b)$ when A is a finite difference approximation to an elliptic operator, and the domain is rectangular. Uses the splitting matrix $M=LL^T$, where $L+L^T$ has the same sparsity structure as A and the coefficients are chosen based on the differential operator. Gives a work estimate of $O(h^{-5/2}\log\varepsilon^{-1})$ in two dimensions to reduce the error by ε, using a Chebyshev sequence of ω's or $O(h^{-3}\log\varepsilon^{-1})$ for certain fixed ω's.

116. /AEL/ Eu, B. E. (1968) "Method of Moments in Collision Theory," *J. Chem. Phys.* 48, pp. 5611-5622.

Uses Lanczos algorithm to compute eigensystem of model of two-body collisions.

117. /PS/ Evans, D. J. (1968) "The Use of Pre-conditioning in Iterative Methods for Solving Linear Equations with Symmetric Positive Definite Matrices," *J. Inst. Maths. Applics.* 4, pp. 295-314.

(Note: The date "1967" on the first page of the article is a misprint.) Considers first-order methods (Gauss-Seidel, etc.) and second-order methods (Richardson, etc.). "Any attempt to improve these basic fundamental methods must clearly apply some form of pre-conditioning to the original equations, in order to minimize the P-condition number and hence increase the rate of convergence." Considers preconditioner $M=(I-\omega L)(I-\omega L^T)$ and applies it to the model problem with ones on the diagonal. Uses Chebyshev acceleration.

118. /C/ Forsythe, George E. (1968) "On the Asymptotic Directions of the s-Dimensional Optimum Gradient Method," *Numer. Math.* 11, pp. 57-76.

Studies the directions from which the iterates approach their limit for the s-dimensional steepest descent algorithm, equivalent to conjugate gradients restarted every s iterations. Shows that either termination occurs in one restart cycle, or convergence is no faster than linear.

119. /ACP/ Fox, R. L. and E. L. Stanton (1968) "Developments in Structural Analysis by Direct Energy Minimization," *AIAA J.* 6, pp. 1036-1042.

Preconditions the stiffness matrix by the diagonal. Reports that conjugate gradients and DFP are then effective.

120. /N/ Horwitz, Lawrence B. and Philip E. Sarachik (1968) "Davidon's Method in Hilbert Space," *SIAM J. Appl. Math.* 16, pp. 676-695.

Uses the same techniques that were applied to the Hestenes-Stiefel (1952) and Fletcher-Reeves (1964) algorithms to generalize algorithm to Hilbert space.

121. /AN/ Klimpel, Richard and Emmet Phillips (1968) "Extrapolation of Thermal Functions to 0^oK Using Unconstrained Nonlinear Optimization," *J. of Chemical and Engineering Data* 13, pp. 97-101.

Uses DFP.

122. /N/ Kratochvil, Alexander (1968) "La Méthode des Gradients Conjugents pour les Equations Non Linéaires das L'Espace de Banach," *Commentationes Mathematicae Universitatis Carolinae* 9, pp. 659-676.

Studies convergence of the conjugate gradient method for monotone nonlinear operators between a reflexive Banach space and its dual.

123. /N/ Lynch, R. T. and K. A. Fegley (1968) "The Davidon Method for Optimal Control Problems," *Proc. IEEE* 56, pp. 1253-1254.

Extends DFP method to finite dimensional function space.

124. /P/ Marchuk, G. I. and Ju. A. Kuznecov (1968) "On Optimal Iteration Processes," *Soviet Math. Dokl.* 9, No. 4, pp. 1041-1045.

(Dokl. Akad. Nauk SSSR 181, No. 6, 1968.) Studies the convergence of an iterative method with polynomial preconditioning, and shows it equivalent to the s-dimensional steepest descent algorithm for certain choices of the coefficients. Discusses the s step conjugate gradient algorithm with preconditioning.

125. /CN/ Myers, Geraldine E. (1968) "Properties of the Conjugate-Gradient and Davidon Methods," *J. of Optimization Theory and Applications* 2, pp. 209-219.

Shows that the algorithms produce directions that are scalar multiples of each other and, under perfect line searches, identical iterates.

126. /AN/ Nagy, George (1968a) "Classification Algorithms in Pattern Recognition," *IEEE Trans. on Audio and Electroacoustics* AU-16, pp. 203-212.

Notes that conjugate gradients can be used in classification according to the Anderson-Bahadur criterion (see Nagy (1968b)).

127. /AN/ Nagy, George (1968b) "State of the Art in Pattern Recognition," *Proc. IEEE* 56, pp. 836-862.

Discusses the use of conjugate gradients in optimizing using the "minimax decision rule" of Anderson and Bahadur, which equalizes the probability of type 1 and type 2 classification errors.

128. /N/ Pagurek, B. and C. M. Woodside (1968) "The Conjugate Gradient Method for Optimal Control Problems with Bounded Variables," *Automatica* 4, pp. 337-349.

Modifies the Fletcher-Reeves (1964) and Lasdon (second derivative) methods to truncate the step in the presence of constraints.

129. /PS/ Stone, Herbert L. (1968) "Iterative Solution of Implicit Approximations of Multidimensional Partial Differential Equations," *SIAM J. Numer. Anal.* 5, pp. 530-558.

Proposes the iteration $Mx_{n+1}=Mx_n-\omega(Ax_n-b)$, where A is a finite difference approximation to an elliptic operator, and the domain is rectangular. Uses the preconditioning matrix $M=LL^T$, where $L+L^T$ has the same sparsity structure as A and the coefficients are chosen based on the differential operator. Proposes cycling from top to bottom and bottom to top on alternate iterations for the application of the preconditioner.

130. /AN/ Wallach, Yehuda (1968) "Gradient Methods for Load-Flow Problems," *IEEE Trans. on Power Apparatus and Systems* PAS-87, pp. 1314-1318.

Formulates load flow problem as an optimization problem and applies steepest descent and conjugate gradients.

131. /L/ Yamamoto, Tetsuro (1968) "On Lanczos' Algorithm for Tri-Diagonalization," *J. Sci. Hiroshima Univ. Ser A-I* 32, pp. 259-284.

Extends the results of Causey and Gregory (1961) on continuing the Lanczos biorthogonal reduction algorithm to the case in which both vectors, not just one, vanish. Gives a geometric interpretation.

1969

132. /AN/ Bierson, B. L. (1969) "A Discrete-Variable Approximation to Optimal Flight Paths," *Astronautica Acta* 14, pp. 157-169.

Uses Fletcher-Reeves (1964) to solve a sequence of unconstrained problems.

133. /AN/ Birta, Louis G. and Peter J. Trushel (1969) "A Comparative Study of Four Implementations of a Dynamic Optimization Scheme," *Simulation* 13, No. 2, pp. 89-97.

Concludes that DFP is faster than Fletcher and Reeves (1964) conjugate gradients on a set of optimal control problems.

134. /CN/ Cantrell, Joel W. (1969) "Relation between the Memory Gradient Method and the Fletcher-Reeves Method," *J. of Optimization Theory and Applications* 4, pp. 67-71.

Notes that memory gradient and Fletcher-Reeves are the same on quadratic functions.

135. /EN/ Fox, R. L. and M. P. Kapoor (1969) "A Minimization Method for the Solution of the Eigenproblem Arising in Structural Dynamics," in *Proceedings of the Second Conference on Matrix Methods in Structural Mechanics*, ed. L. Berke, R. M. Bader, W. J. Mykytow, J. S. Przemieniecki, M. H. Shirk, Wright-Patterson Air Force Base, Ohio AFFDL-TR-68-150, pp. 271-306.

Finds several small eigenvalues of a generalized eigenproblem by using the Bradbury-Fletcher (1966) idea of minimizing the Raleigh quotient using the conjugate gradient algorithm, and the idea of orthogonalizing against the eigenvectors previously determined. Gives numerical examples.

136. /AC/ Fried, Isaac (1969) "More on Gradient Iterative Methods in Finite-Element Analysis," *AIAA J.* 7, pp. 565-567.

Uses conjugate gradients to construct the explicit inverse of a finite-element matrix, and discusses storage management on tape units, keeping the original matrix unassembled. Discusses modifications to the matrix in case it is rank deficient.

137. /AEL/ Garibotti, C. R. and M. Villani (1969) '"Continuation in the Coupling Constant for the Total *K* and *T* Matrices," *Il Nuovo Cimento* 59, pp. 107-123.

Uses the Lanczos (1950) algorithm for finding the eigensystem of a problem in nonrelativistic scattering theory.

138. /CP/ Godunov, S. K. and G. P. Prokopov (1969) "Solution of the Laplace Difference Equation," *USSR Comp. Math. and Math. Phys.* 9, No. 2, pp. 285-292.

(Zh. vўchisl. Mat. mat. Fiz. 9, No. 2, pp. 462-468.) For the model problem, obtains an algorithm with number of iterations independent of mesh size by combining ADI with a Rayleigh-Ritz criteria for the parameters.

139. /CN/ Goldfarb, Donald (1969) "Extension of Davidon's Variable Metric Method to Maximization under Linear Inequality and Equality Constraints," *SIAM J. Appl. Math.* 17, pp. 739-764.

Proposes a stable gradient projection approach updating the full Hessian approximation.

140. /N/ Hestenes, Magnus R. (1969) "Multiplier and Gradient Methods," *J. of Optimization Theory and Applications* 4, pp. 303-320.

Discusses conjugate direction and conjugate gradients in terms of the Rayleigh-Ritz method on Hilbert space with applications to minimizing a function subject to nonlinear constraints by the augmented Lagrangian method.

141. /N/ Kawamura, K. and R. A. Volz (1969) "On the Convergence of the Conjugate Gradient Method in Hilbert Space," *IEEE Trans. on Auto. Control* AC-14, pp. 296-297.

Extends conjugate gradients to Hilbert spaces using uniform continuity of the gradient instead of bounded second Frechet derivatives as do Mitter, Lasdon, and Waren (1966).

142. /AC/ Luenberger, David G. (1969) "Hyperbolic Pairs in the Method of Conjugate Gradients," *SIAM J. Appl. Math.* 17, pp. 1263-1267.

Applies conjugate gradients to the indefinite matrix system corresponding to minimizing a quadratic subject to linear equality constraints. Derives a double step algorithm to overcome breakdown when the direction vector satisfies $(p,Ap)=0$.

143. /AL/ Marshall, Jr., Thomas G. (1969) "Synthesis of RLC Ladder Networks by Matrix Tridiagonalization," *IEEE Trans. Circuit Theory* CT-16, pp. 39-46.

Reduces cyclic matrices to tridiagonal form by an algorithm which has the Lanczos method as a special case.

144. /CN/ Mehra, Raman K. (1969) "Computation of the Inverse Hessian Matrix Using Conjugate Gradient Methods," *Proc. IEEE* 57, pp. 225-226.

Constructs the Hessian inverse from the search directions.

145. /N/ Miele, A., H. Y. Huang, and J. C. Heideman (1969) "Sequential Gradient-Restoration Algorithm for the Minimization of Constrained Functions -- Ordinary and Conjugate Gradient Versions," *J. of Optimization Theory and Applications* 4, pp. 213-243.

Gives an extension of the conjugate gradient algorithm to minimization of nonlinear functions subject to nonlinear equality constraints by alternating conjugate gradient steps on the augmented Lagrangian function with steps back to the constraints.

146. /ACN/ Pearson, J. D. (1969) "Variable Metric Methods of Minimisation," *Computer J.* 12, pp. 171-178.

Concludes that conjugate gradients is generally better than DFP for well-conditioned problems and worse for ill-conditioned.

147. /N/ Polak, E. and G. Ribiere (1969) "Note sur la Convergence de Methodes de Directions Conjugees," *Revue Francaise d'Informatique et de Recherche Operationnelle* 3, pp. 35-43.

Modifies the update to the direction vector in the Fletcher-Reeves (1964) algorithm.

148. /CN/ Polyak, B. T. (1969) "The Conjugate Gradient Method in Extremal Problems," *USSR Comp. Math. and Math. Phys.* 9, No. 4, pp. 94-112.

(Zh. vy̌chisl. Mat. mat. Fiz. 9, No. 4, pp. 807-821.) Proves convergence of a conjugate gradient method for nonquadratic functions and for quadratic functions with upper and lower bounds on the variables. Advocates saving the direction vectors and using them for a change of basis.

149. /AEL/ Sebe, T. and J. Nachamkin (1969) "Variational Buildup of Nuclear Shell Model Bases," *Annals of Physics* 51, pp. 100-123.

Uses the Lanczos algorithm with very few steps to find eigenvalues corresponding to low lying states belonging to nuclear spins in a shell model.

150. /CN/ Sorenson, H. W. (1969) "Comparison of Some Conjugate Direction Procedures for Function Minimization," *J. of the Franklin Institute* 288, pp. 421-441.

Shows that DFP, conjugate gradients, and Partan are identical on quadratic functions. Derives other properties. Proposes a new definition for the β parameter: $\Delta g^T g / \Delta g^T p$.

151. /AN/ Westcott, J. H. (1969) "Numerical Computational Methods of Optimisation in Control," *Automatica* 5, pp. 831-843.

Gives survey of methods, including conjugate gradients.

1970

152. /CN/ Broyden, C. G. (1970) "The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations," *J. Inst. Maths. Applics.* 6, pp. 76-90.

Analyzes error vectors $x_k - x^*$ in the Broyden (1967) family of Quasi-Newton algorithms when applied to the minimization of a quadratic function. Uses the observations to gain insight into the algorithms' behavior on non-quadratic functions.

153. /AL/ Chang, F. Y. and Omar Wing (1970) "Multilayer *RC* Distributed Networks," *IEEE Trans. on Circuit Theory* CT-17, pp. 32-40.

Uses Lanczos algorithm to generate a tridiagonal matrix with positive entries, related to the physical parameters in the required network.

154. /CN/ Daniel, James W. (1970a) *The Approximate Minimization of Functionals,* Prentice-Hall, Englewood Cliffs, New Jersey.

Discusses the conjugate gradient algorithm in Hilbert space, including conjugate directions on quadratics, conjugate gradients on quadratics, and general conjugate gradients.

155. /CN/ Daniel, James W. (1970b) "A Correction Concerning the Convergence Rate for the Conjugate Gradient Method," *SIAM J. Numer. Anal.* 7, pp. 277-280.

Gives correction to Daniel (1967) result for nonlinear equations.

156. /AC/ De, S. and A. C. Davies (1970) "Convergence of Adaptive Equaliser for Data Transmission," *Electronics Letters* 6, pp. 858-861.

Proposes conjugate gradients for solving a least squares problem.

157. /AL/ Emilia, David A. and Gunnar Bodvarsson (1970) "More on the Direct Interpretation of Magnetic Anomalies," *Earth and Planetary Science Letters* 8, pp. 320-321.

Relates the convergence theory for their 1969 algorithm to that of the Lanczos minimized iterations algorithm.

158. /AC/ George, J. Alan (1970) *The Use of Direct Methods for the Solution of the Discrete Poisson Equation on Non-Rectangular Regions,* STAN-CS-70-159, Computer Science Department, Stanford University, Stanford, California.

Suggests solving the Poisson equation on domains that are unions or differences of rectangles by taking advantage of fast solvers for rectangular regions. Suggests using a capacitance matrix method of Hockney or applying the Sherman-Morrison-Woodbury formula, recognizing that the desired equation involves a matrix which is a low rank modification of one that can be handled by fast solvers. Proposes solving the resulting systems by direct methods or iterative methods. Proposes SOR or conjugate gradients for the capacitance matrix algorithm, since the matrix may not be explicitly available. Proposes a termination criterion for the iterative methods. Presents numerical experiments using the SOR algorithm.

159. /EL/ Godunov, S. K. and G. P. Prokopov (1970) "A Method of Minimal Iterations for Evaluating the Eigenvalues of an Elliptic Operator," *USSR Comp. Math. and Math. Phys.* 10, No. 5, pp. 141-154.

(Zh. vÿchisl. Mat. mat. Fiz. 10, No. 5, pp. 1180-1190.) Uses the Lanczos (1950) algorithm to reduce the matrix of a difference operator to tridiagonal form, and computes the frequencies of a piezo-electric resonator. Recomputes the eigenvalues for increasing number of Lanczos steps until there is little change. Notes that taking more than n steps causes extra copies of the eigenvalues to appear and converge.

160. /CN/ Huang, H. Y. (1970) "Unified Approach to Quadratically Convergent Algorithms for Function Minimization," *J. of Optimization Theory and Applications* 5, pp. 405-423.

Derives a class of algorithms with one-dimensional line searches, quadratic termination, function and gradient evaluations only, and using only that information available at current and previous step. Notes that variable metric algorithms and conjugate gradient algorithms are special cases.

161. /ACP/ Kamoshida, Mototaka, Kenji Kani, Kozo Sato, and Takashi Okada (1970) "Heat Transfer Analysis of Beam-Lead Transistor Chip," *IEEE Trans. on Electron. Devices* ED-17, pp. 863-870.

Uses conjugate gradients with scaling to make the diagonal of the matrix equal to 1.

162. /N/ Kelley, H. J. and J. L. Speyer (1970) "Accelerated Gradient Projection," in *Symposium on Optimization*, ed. A. V. Balakrishnan, M. Contensori, B. F. de Veubeke, P. Kree, J. L. Lions and N. N. Moiseev, Lecture Notes in Mathematics 132, Springer, New York, pp. 151-158.

Develops the DFP algorithm for nonlinear constraints.

163. /AC/ Kobayashi, Hisashi (1970) "Iterative Synthesis Methods for a Seismic Array Processor," *IEEE Trans. on Geoscience Electronics* GE-8, pp. 169-178.

Uses a few steps of the conjugate gradient algorithm with projection to minimize a quadratic function subject to a constraint.

164. /ACP/ Luenberger, David G. (1970) "The Conjugate Residual Method for Constrained Minimization Problems," *SIAM J. Numer. Anal.* 7, pp. 390-398.

Constructs a method with residuals A-conjugate and directions A^2-conjugate and applies it to quadratic minimization with linear equality constraints.

165. /N/ Miele, A. and J. W. Cantrell (1970) "Memory Gradient Method for the Minimization of Functions," in *Symposium on Optimization*, ed. A. V. Balakrishnan, M. Contensori, B. F. de Veubeke, P. Kree, J. L. Lions and N. N. Moiseev, Lecture Notes in Mathematics 132, Springer, New York, pp. 252-263.

Develops a method for nonquadratic functions which takes steps of the form $-\alpha g+\beta p_{old}$, minimizing at each step over α and β. Reduces to the conjugate gradient method in the case of quadratic functions.

166. /ACE/ Ojalvo, I. U. and M. Newman (1970) "Vibration Modes of Large Structures by an Automatic Matrix-Reduction Method," *AIAA J.* 8, pp. 1234-1239.

Solves the generalized eigenvalue problem with the Lanczos algorithm using a small number of steps (credited to Crandall (1956)) and the Causey-Gregory (1961) stabilization method.

167. /N/ Ortega, James M. and Werner C. Rheinboldt (1970) "Local and Global Convergence of Generalized Linear Iterations," in *Studies in Numerical Analysis 2: Numerical Solutions of Nonlinear Problems*, ed. J. M. Ortega and W. C. Rheinboldt, SIAM, Philadelphia.

Gives convergence results for conjugate gradients under conditions such as that the function is twice continuously differentiable and smallest eigenvalue of the Hessian is bounded below by a number greater than zero on all of R^n.

168. /EL/ Paige, C. C. (1970) "Practical Use of the Symmetric Lanczos Process with Re-Orthogonalization," *BIT* 10, pp. 183-195.

Gives rounding error analysis and stopping criterion for finding several extreme eigenvalues and corresponding eigenvectors using complete re-orthogonalization. Notes that eigenestimates may be accurate despite loss of orthogonality.

169. /AEL/ Peters, G. and J. H. Wilkinson (1970) "$Ax=\lambda Bx$ and the Generalized Eigenproblem," *SIAM J. Numer. Anal.* 7, pp. 479-492.

Applies Lanczos algorithm with complete reorthogonalization to $L^{-1}AL^{-T}$ where $B=LL^T$. Discusses Golub's idea for complete reorthogonalization without saving the vectors.

170. /N/ Powell, M. J. D. (1970) "A Survey of Numerical Methods for Unconstrained Optimization," *SIAM Rev.* 12, pp. 79-97.

Surveys nonlinear conjugate gradients and conjugate direction methods, including s-step gradient, Zoutendijk's (1960) method, Fletcher-Reeves (1964), Powell, and DFP.

171. /AN/ Smith, Otto J. M (1970) "Power System State Estimation," *IEEE Trans. on Power Apparatus and Systems* PAS-89, pp. 375-380.

Uses DFP in solving a least squares problem related to power distribution.

172. /AN/ Straeter, Terry A. and John E. Hogge (1970) "A Comparison of Gradient Dependent Techniques for the Minimization of an Unconstrained Function of Several Variables," *AIAA J.* 8, pp. 2226-2229.

Compares DFP and Fletcher-Reeves (1964) algorithm to other methods on problems related to optimal control.

173. /N/ Tripathi, S. S. and K. S. Narendra (1970) "Optimization Using Conjugate Gradient Methods," *IEEE Trans. on Auto. Control* AC-15, pp. 268-269.

Uses DFP on optimal control problems. Contrasts with Lynch and Fegley (1968) in that the method is applied to the original problem, not a discretized version.

1971

174. /ACN/ Fong, T. S. and R. A. Birgenheier (1971) "Method of Conjugate Gradients for Antenna Pattern Synthesis," *Radio Science* 6, pp. 1123-1130.

Uses conjugate gradients to minimize an error function with Frechet derivative.

175. /AENP/ Geradin, M. (1971) "The Computational Efficiency of a New Minimization Algorithm for Eigenvalue Analysis," *J. of Sound and Vibration* 19, pp. 319-331.

Uses diagonal scaling and the local Hessian in the computation of conjugate gradient parameters for minimizing the Rayleigh quotient for eigenvalues of a plate problem.

176. /AN/ Goldberg, Saul and Allen Durling (1971) "A Computational Algorithm for the Identification of Nonlinear Systems," *J. of the Franklin Institute* 291, pp. 427-447.

Solves a nonlinear control problem using conjugate gradients.

177. /N/ Kelley, H. J. and G. E. Myers (1971) "Conjugate Direction Methods for Parameter Optimization," *Astronautica Acta* 16, pp. 45-51.

Compares five methods and finds conjugate gradients better than Davidon.

178. /AN/ Kobayashi, Hisashi (1971) "Simultaneous Adaptive Estimation and Decision Algorithm for Carrier Modulated Data Transmission Systems," *IEEE Trans. on Commun. Tech.* COM-19, pp. 268-280.

Proposes solving maximum likelihood problems by conjugate gradients.

179. /C/ Maistrovskii, G. D. (1971) "Convergence of the Method of Conjugate Gradients," *USSR Comp. Math. and Math. Phys.* 11 No. 5, pp. 244-248.

(Zh. vy̌chisl. Mat. mat. Fiz. 11, No. 5, pp. 1291-1294.) Proves that the Fletcher-Reeves (1964) algorithm converges for any uniformly convex function with bounded level sets and Lipschitz continuous gradient, using exact line searches.

180. /EL/ Paige, C. C. (1971) *The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices,* Ph. D. dissertation, University of London.

Gives rounding error analysis of Lanczos (1950) method (and Hessenberg methods in general). Corrects and expands Kaniel (1966) convergence theory. Compares various implementations.

181. /AL/ Phillips, James L. (1971) "The Triangular Decomposition of Hankel Matrices," *Math. of Comp.* 25, pp. 599-602.

Observes that a Hankel matrix is a moment matrix: $H_{ij}=(B^{i-1}v,B^{j-1}v)$ for some matrix B and vector v. Applies the Lanczos (1950) algorithm to B and obtains a Cholesky factorization of H as a byproduct of $O\,(n^2)$ operations.

182. /CN/ Powell, M. J. D. (1971) "Recent Advances in Unconstrained Optimization," *Math. Programming* 1, pp. 26-57.

Surveys conjugate gradients and Quasi-Newton research (among other things) from 1967 to 1971.

183. /C/ Reid, J. K. (1971) "On the Method of Conjugate Gradients for the Solution of Large Sparse Systems of Linear Equations," in *Large Sparse Sets of Linear Equations*, Academic Press, New York, pp. 231-254.

Emphasizes the use of conjugate gradients as an iterative algorithm for large, sparse, well-conditioned problems, using much fewer than n iterations. Discusses various computational forms of the algorithm and compares storage requirements, operations counts, and stability. Recommends recursive residuals with $\alpha=r^T r/p^T Ap$ and $\beta=r^T r/r^T r$. Also uses minimum residual algorithm.

184. /EL/ Takahasi, Hidetosi and Makoto Natori (1971-72) "Eigenvalue Problem of Large Sparse Matrices," *Rep. Compt. Centre, Univ. Tokyo* 4, pp. 129-148.

Performs a stability analysis for the Lanczos (1950) algorithm without reorthogonalization, and proposes stopping the iteration when the inner product of the Lanczos iterate with the initial vector grows too large in order to prevent round-off errors in the estimated eigenvalues. Gives numerical experiments illustrating the effects of the stopping criterion.

185. /EL/ Weaver, Jr., William and David M. Yoshida (1971) "The Eigenvalue Problem for Banded Matrices," *Computers and Structures* 1, pp. 651-664.

Uses Lanczos (1950) algorithm with n iterations and full reorthogonalization to solve the generalized banded eigenvalue problem $Ax=\lambda Bx$ where B is symmetric and positive definite and A is symmetric. Solves linear systems involving B and uses QR on the tridiagonal matrix.

186. /PS/ Widlund, Olof B. (1971) "On the Effects of Scaling of the Peaceman-Rachford Method," *Math. of Comp.* 25, pp. 33-41.

Analyzes and discusses the choice of D in the iteration $(\omega D^2+H)x_{n+1/2}=(\omega D^2-V)x_n+b$, $(\omega D^2+V)x_{n+1}=(\omega D^2-H)x_{n+1/2}+b$, and suggests the use of $D=diag\,(H)$ or $D=diag\,(V)$.

187. /AN/ Willoughby, J. K. and B. L. Pierson (1971) "A Constraint-Space Conjugate Gradient Method for Function Minimization and Optimal Control Problems," *Int. J. Control* 14, pp. 1121-1135.

Applies conjugate gradients to the Lagrangian function, using line search to guarantee satisfaction of equality constraints.

1972

188. /ACPS/ Axelsson, O. (1972) "A Generalized SSOR Method," *BIT* 12, pp. 443-467.

Solves a nonseparable elliptic partial differential equation by conjugate gradients preconditioned with a scaled SSOR operator based on a mesh varying with the smoothness of the coefficients. Shows that the number of iterations is dependent on $h^{-1/2}$ when h is defined by $h^2\lambda$ being the smallest eigenvalue of $D^{-1}A$ as the mesh-size goes to zero. Presents numerical experiments.

189. /CN/ Beale, E. M. L. (1972) "A Derivation of Conjugate Gradients," in *Numerical Methods for Non-linear Optimization*, ed. F. A. Lootsma, Academic Press, New York, pp. 39-43.

Gives an elementary derivation of the algorithm, including the case where the initial direction is not the gradient.

190. /CN/ Broyden, C. G. and M. P. Johnson (1972) "A Class of Rank-1 Optimization Algorithms," in *Numerical Methods for Non-linear Optimization*, ed. F. A. Lootsma, Academic Press, New York, pp. 35-38.

Derives an update formula based on minimizing a norm of the difference between the inverse Jacobian and the approximation matrix.

191. /CN/ Cohen, Arthur I. (1972) "Rate of Convergence of Several Conjugate Gradient Algorithms," *SIAM J. Numer. Anal.* 9, pp. 248-259.

Proves an n-step quadratic convergence rate for the Polak-Ribiere (1969), Daniel (1967), and Fletcher-Reeves (1964) algorithms when they are reinitialized periodically. Corrects errors in the work of Daniel and of Polyak (1969).

192. /N/ Crowder, Harlan and Philip Wolfe (1972) "Linear Convergence of the Conjugate Gradient Method," *IBM J. of Res. and Devel.* 16, pp. 431-433.

Shows that conjugate gradients with initial direction vector not equal to the residual converges only linearly, and that conjugate gradients with no restarts on a nonlinear function converges no worse than linearly.

193. /AEL/ Dahlquist, Germund, Stanley C. Eisenstat, and Gene H. Golub (1972) "Bounds for the Error of Linear Systems of Equations Using the Theory of Moments," *J. of Math. Anal. and Applics.* 37, pp. 151-166.

Uses the Lanczos algorithm to derive quadrature rules by finding roots of a set of orthogonal polynomials, and uses these for obtaining error bounds for linear systems.

194. /CN/ Dennis, Jr., J. E. (1972) "On Some Methods Based on Broyden's Secant Approximation to the Hessian," in *Numerical Methods for Non-linear Optimization*, ed. F. A. Lootsma, Academic Press, New York, pp. 19-34.

Surveys some rank-one and rank-two update algorithms and a class of methods for least squares problems and some convergence results.

195. /C/ Devooght, J. (1972) "The Reproducing Kernel Method II," *J. Math. Phys.* 13, pp. 1259-1268.

Derives a connection between the conjugate gradient algorithm and the reproducing kernel method.

196. /CN/ Dixon, L. C. W. (1972) "Quasi-Newton Algorithms Generate Identical Points," *Math. Programming* 2, pp. 383-387.

Notes that all algorithms in the Broyden (1967) family of algorithms generate the same sequence of points on general differentiable functions if the line searches are exact.

197. /AEN/ Fried, I. (1972) "Optimal Gradient Minimization Scheme for Finite Element Eigenproblems," *J. of Sound and Vibration* 20, pp. 333-342.

Computes the smallest eigenvalue of a generalized eigenvalue problem by applying conjugate gradient to minimize the Rayleigh quotient. Determines the β parameter by seeking the direction which will result in the lowest function value, much as in the memory gradient method. Uses projection to get the higher eigenvalues. Estimates the error in the eigenvalue estimates and presents numerical experiments.

198. /AEL/ Golub, G. H., R. Underwood, and J. H. Wilkinson (1972) *The Lanczos Algorithm for the Symmetric $Ax=\lambda Bx$ Problem,* STAN-CS-72-270, Stanford University Computer Science Department Report, Stanford, California.

Assumes B positive definite, and iterates using a Cholesky factorization of it. Gives an Algolw program.

199. /L/ Gragg, W. B. (1972) "The Padé Table and Its Relation to Certain Algorithms of Numerical Analysis," *SIAM Rev.* 14, pp. 1-62.

Gives the relation between the Lanczos polynomials and the Pade Table.

200. /EL/ Haydock, R., Volker Heine, and M. J. Kelley (1972) "Electronic Structure Based on the Local Atomic Environment for Tight-Binding Bands," *J. Phys. C: Solid State Physics* 5, pp. 2845-2858.

Independently discovers the Lanczos (1950) algorithm for symmetric matrices and names it the "recursion method." Does not normalize the vectors.

201. /C/ Kammerer, W. J. and M. Z. Nashed (1972a) "Iterative Methods for Best Approximate Solutions of Linear Integral Equations of the First and Second Kinds," *J. of Math. Anal. and Applics.* 40, pp. 547-573.

Proves that conjugate gradients converges to a least squares solution, and under certain conditions, to the one of minimal norm.

202. /C/ Kammerer, W. J. and M. Z. Nashed (1972b) "On the Convergence of the Conjugate Gradient Method for Singular Linear Operator Equations," *SIAM J. Numer. Anal.* 9, pp. 165-181.

Proves that conjugate gradients, applied to minimizing the norm of the residual of an equation involving a bounded linear operator between two Hilbert spaces with closed range converges to a least squares solution. Gives bounds on rate of convergence in both cases. Also studies the case of nonclosed range.

203. /N/ Klessig, R. and E. Polak (1972) "Efficient Implementations of the Polak-Ribiere Conjugate Gradient Algorithm," *SIAM J. Control* 10, pp. 524-549.

Presents two modifications to the conjugate gradient algorithm to ensure convergence without line searches.

204. /AN/ Leondes, C. T. and C. A. Wu (1972) "The Conjugate Gradient Method and Its Application to Aerospace Vehicle Guidance and Control. Part I: Basic Results in the Conjugate Gradient Method. Part II: Mars Entry Guidance and Control," *Astronautica Acta* 17, pp. 871-890.

Calculates optimal trajectory for simulated problem of braking into Mars atmosphere.

205. /CN/ McCormick, Garth P. and Klaus Ritter (1972) "Methods of Conjugate Directions vs. Quasi-Newton Methods," *Math. Programming* 3, pp. 101-116.

Recommends the one-step superlinear Quasi-Newton algorithms over the n or $n-1$-step superlinear conjugate direction methods.

206. /EL/ Paige, C. C. (1972) "Computational Variants of the Lanczos Method for the Eigenproblem," *J. Inst. Maths. Applics.* 10, pp. 373-381.

Gives round-off error analysis and computational experience with various mathematically equivalent formulations.

207. /N/ Pierson, B. L. (1972) "A Modified Conjugate Gradient Method for Optimization Problems," *Int. J. Control* 16, pp. 1193-1196.

Shows experimentally that restarting a nonlinear conjugate gradient procedure is as good as the Mehra (1969) idea of taking a Newton step every n iterations, based on the accumulation of the Hessian by the conjugate gradient directions.

208. /CN/ Powell, M. J. D. (1972) "Some Properties of the Variable Metric Algorithm," in *Numerical Methods for Non-linear Optimization*, ed. F. A. Lootsma, Academic Press, New York, pp. 1-17.

Proves convergence of the DFP algorithm in case that a level set is bounded and the function is convex (not uniformly convex).

209. /AC/ Reid, J. K. (1972) "The Use of Conjugate Gradients for Systems of Linear Equations Possessing 'Property A'," *SIAM J. Numer. Anal.* 9, pp. 325-332.

Iterates on half of the variables to save work.

210. /AC/ Ruhe, Axel and Torbjörn Wiberg (1972) "The Method of Conjugate Gradients Used in Inverse Iteration," *BIT* 12, pp. 543-554.

Shows that Golub's idea for solving the inverse iteration equation by conjugate gradients often gives convergence in a small number of steps.

211. /ACN/ Takahashi, Tomowaki (1972) "An Experimental Analysis of Optimization Algorithms Using a Model Function," *Optik* 35, pp. 101-115.

Recommends against conjugate gradients for nonlinear problems.

212. /AEL/ Whitehead, R. R. (1972) "A Numerical Approach to Nuclear Shell-Model Calculations," *Nuclear Physics* A182, pp. 290-300.

Uses the Lanczos (1950) algorithm to compute approximate eigenstates.

213. /PS/ Young, David M. (1972) "Second-Degree Iterative Methods for the Solution of Large Linear Systems," *J. Approx. Theory* 5, pp. 137-148.

Notes that second-order Richardson is an acceleration of Jacobi. "However, it does not seem to be generally recognized that second degree methods can be effectively applied to other methods as well." Estimates rates of convergence for acceleration of SSOR, improving the Habetler and Wachspress (1961) method of estimating ω.

1973

214. /AN/ Bloemer, William L. and Buddy L. Bruner (1973) "Optimization of Variational Trial Functions," *J. Chem. Phys.* 58, pp. 3735-3744.

Uses DFP algorithm for computations in atomic and molecular theory.

215. /S/ Bracha-Barak, Amnon and Paul E. Saylor (1973) "A Symmetric Factorization Procedure for the Solution of Elliptic Boundary Value Problems," *SIAM J. Numer. Anal.* 10, pp. 190-206.

Studies an algorithm proposed by Stone (private communication, 1969) which, although not second order (see Saylor (1974)), is designed to force maximal cancellation when the error matrix is applied to a discretization of a first-degree polynomial.

216. /PS/ Concus, Paul and Gene H. Golub (1973) "Use of Fast Direct Methods for the Efficient Numerical Solution of Nonseparable Elliptic Equations," *SIAM J. Numer. Anal.* 10, pp. 1103-1120.

Uses a preconditioned Chebyshev iteration, with fast direct solution of Helmholtz's equation on a rectangle, to solve linear and self-adjoint elliptic partial differential equations. Obtains convergence estimates independent of mesh size for scaled equations that are smooth. Proposes fast direct solution of general separable operators as a preconditioner.

217. /CPS/ Evans, D. J. (1973) "The Analysis and Application of Sparse Matrix Algorithms in the Finite Element Method," in *The Mathematics of Finite Elements and Applications*, ed. J. R. Whiteman, Academic Press, New York, pp. 427-447.

Surveys direct methods, SOR variants, preconditioned Richardson and Chebyshev methods, gradient methods, and preconditioned conjugate gradient. Gives results for a model biharmonic problem using conjugate gradients with SSOR preconditioning and various relaxation parameters.

218. /AEL/ Golub, G. H. (1973) "Some Uses of the Lanczos Algorithm in Numerical Linear Algebra," in *Topics in Numerical Analysis*, ed. John J. H. Miller, Academic Press, New York, pp. 173-184.

Gives clear derivation of the Lanczos (1950) algorithm for the symmetric eigenvalue problem and for the solution of linear systems. Gives error bounds on the eigenvalues and on the errors in the linear system after fewer than n steps.

219. /C/ Kielbasiński, A., Grazyna Woźniakowska, and H. Woźniakowski (1973) "Algorytmizacja metod najlepszej strategii dla wielkich ukladów równań o symetrycznej, dodatnio określonej macierzy," *Roczniki Polskiego Towarzystwa Matematycznego: Matematyka Stosowana* Seria 3,1, pp. 47-68.

"Algorithmization of the Best Strategy Methods for Large Linear Systems with a Positive Definite Matrix" presents an Algol code for a method which combines the Chebyshev and minimal residual iterations (see Engeli et al. (1959), Chapter 2).

220. /CN/ Luenberger, David G. (1973) *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Menlo Park, California.

Gives derivation of the conjugate gradient method from the viewpoint of conjugate directions and optimal polynomials. Develops the partial conjugate gradient algorithm (see also Forsythe (1968)), the expanding subspace property, and convergence bounds, and extends the algorithm to nonquadratic problems. Summarizes Luenberger's research in the field.

221. /AN/ Polak, E. (1973) "An Historical Survey of Computational Methods in Optimal Control," *SIAM Rev.* 15, pp. 553-584.

Surveys gradient and other methods.

222. /ACN/ Powers, W. F. (1973) "A Crude-Search Davidon-Type Technique with Applications to Shuttle Optimization," *J. Spacecraft* 10, pp. 710-715.

Uses DFP with crude line searches.

223. /CP/ Stewart, G. W. (1973) "Conjugate Direction Methods for Solving Systems of Linear Equations," *Numer. Math.* 21, pp. 285-297.

Develops bi-conjugate direction methods, using one set of basis vectors for the search directions and another to define a subspace orthogonal to the residual. Relates the algorithms to matrix factorizations. Develops the block generalization of this class of algorithms.

224. /AC/ Wang, R. J. and S. Treitel (1973) "The Determination of Digital Wiener Filters by Means of Gradient Methods," *Geophysics* 38, pp. 310-326.

Analyzes seismic data by least squares and conjugate gradients.

1974

225. /ACP/ Axelsson, O. (1974a) *On Preconditioning and Convergence Acceleration in Sparse Matrix Problems,* CERN Technical Report 74-10, Data Handling Division, Geneva.

Proposes conjugate gradients or Chebyshev iteration preconditioned by SSOR (or block-SSOR) for solving discretizations of elliptic partial differential equations. Proves that, for the model problem, the condition number is reduced by the power ½. Gives numerical results for the model problem.

226. /ACP/ Axelsson, O. (1974b) "On the Efficiency of a Class of A-Stable Methods," *BIT* 14, pp. 279-287.

Uses the term "preconditioning." Solves a particular linear system by preconditioning conjugate gradients or Chebyshev with a related linear operator, giving condition number independent of mesh.

227. /ACNP/ Bartels, Richard and James W. Daniel (1974) "A Conjugate Gradient Approach to Nonlinear Elliptic Boundary Value Problems in Irregular Regions," in *Conference on the Numerical Solution of Differential Equations, Dundee, 1973,* ed. G. A. Watson, Springer Verlag, New York.

Develops the idea in Daniel's 1965 thesis of solving discretizations of linear or nonlinear self-adjoint elliptic partial differential equations by conjugate gradients, preconditioned by the Laplacian operator. Uses a fast Poisson solver at each iteration. Shows that the convergence rate is independent of mesh size. Provides numerical results.

228. /CEL/ Belford, Geneva G. and E. H. Kaufman, Jr. (1974) "An Application of Approximation Theory to an Error Estimate in Linear Algebra," *Math. of Comp.* 28, pp. 711-712.

Corrects a result of Kaniel (1966) on the convergence of the conjugate gradient and Lanczos algorithms by noting that the standard Chebyshev theorem does not apply although a result of Kaufman and Belford *(J. Approx. Theory 7 (1973), pp. 21-35)* gives the desired conclusion.

229. /CN/ Bertsekas, Dimitri P. (1974) "Partial Conjugate Gradient Methods for a Class of Optimal Control Problems," *IEEE Trans. on Auto. Control* AC-19, pp. 209-217.

Uses conjugate gradients preconditioned by the inverse of a part of the Hessian evaluated at the initial point. Restarts the iteration every $s<n$ steps.

230. /S/ Bracha-Barak, Amnon (1974) "A Factorization Procedure for the Solution of Multidimensional Elliptic Partial Differential Equations," *SIAM J. Numer. Anal.* 11, pp. 887-893.

Generalizes the Stone symmetric splitting of Bracha-Barak and Saylor (1973) to more than two space dimensions and studies convergence properties.

231. /EL/ Cullum, Jane and W. E. Donath (1974a) *A Block Generalization of the Symmetric S-Step Lanczos Algorithm,* IBM T. J. Watson Research Center Report RC 4845, Yorktown Heights, New York.

Develops the block Lanczos algorithm with selected reorthogonalization. Presents numerical results.

232. /EL/ Cullum, Jane and W. E. Donath (1974b) "A Block Lanczos Algorithm for Computing the q Algebraically Largest Eigenvalues and a Corresponding Eigenspace for Large, Sparse Symmetric Matrices," in *Proc. 1974 IEEE Conference on Decision and Control*, IEEE Press, New York, pp. 505-509.

Presents the block Lanczos algorithm as a generalization of the algorithm of Karush (1951).

233. /PS/ Evans, D. J. (1974) "Iterative Sparse Matrix Algorithms," in *Software for Numerical Mathematics*, ed. D. J. Evans, Academic Press, New York, pp. 49-83.

Surveys the basic iterative methods and discusses the importance of preconditioning in the stationary iterative methods as well as higher order ones such as Richardson extrapolation for large sparse, as well as small dense ill-conditioned matrices. Draws attention to the Lanczos (1950) method as a promising one for computing eigenvalues.

234. /AEL/ Harms, Edward (1974) "A Modified Method of Moments Approach to the Solution of Scattering Equations," *Nuclear Physics* A222, pp. 125-139.

Uses the Lanczos (1950) algorithm for finding eigenvalues relating to scattering theory.

235. /AN/ Haschemeyer, Rudy H. and Leonard F. Estis (1974) "Analysis of Self-Associating Systems from Sedimentation Velocity Data," *J. Biological Chem.* 249, pp. 489-491.

Uses DFP.

236. /AENP/ Hasselman, T. K. and Gary C. Hart (1974) "A Minimization Method for Treating Convergence in Modal Synthesis," *AIAA J.* 12, pp. 316-323.

Minimizes the Rayleigh quotient using conjugate gradients preconditioned by a diagonal matrix. Projects against earlier eigenvectors.

237. /CN/ Huang, H. Y. (1974) "Method of Dual Matrices for Function Minimization," *J. of Optimization Theory and Applications* 13, pp. 519-537.

Presents a method based on two matrices: one for generating (conjugate) directions and the other to generate a descent direction. Proves $n+1$ or fewer step termination on quadratic functions, and needs no line searches.

238. /PS/ Kincaid, David R. (1974) "On Complex Second-Degree Iterative Methods," *SIAM J. Numer. Anal.* 11, pp. 211-218.

Discusses acceleration of stationary iterative methods when the eigenvalues are contained within an ellipse in the complex plane. Notes that the method does not have practical advantage for SOR, but does produce a more efficient algorithm using Gauss-Seidel.

239. /CN/ McCormick, G. P. and K. Ritter (1974) "Alternate Proofs of the Convergence Properties of the Conjugate-Gradient Method," *J. of Optimization Theory and Applications* 13, pp. 497-518.

Proves superlinear convergence of a reset conjugate gradient algorithm, similar to the Polak-Ribiere algorithm, with approximate line searches, compact level sets, f twice continuously differentiable in a neighborhood of x^*, and $f''(x^*)$ positive definite. Gives a rate of convergence when f'' is Lipschitz in a neighborhood of x^*.

240. /EL/ Paige, C. C. (1974) "Bidiagonalization of Matrices and Solution of Linear Equations," *SIAM J. Numer. Anal.* 11, pp. 197-209.

Shows that the Golub and Kahan (1965) algorithm is equivalent to the Lanczos (1950) algorithm applied to a matrix with 0 blocks on the diagonal and A and A^H off the diagonal. Applies the algorithm to solving linear least squares problems and computing eigenvalues of 2-cyclic matrices.

241. /ACP/ Palmer, John F. (1974) *Conjugate Direction Methods and Parallel Computing,* Ph.D. dissertation, Stanford University Computer Science Department, Stanford, California.

Develops the block form of the Golub and Kahan (1965) bidiagonalization algorithm. Adds reorthogonalization to Luenberger (1973) partial conjugate gradient algorithm. Derives block Lanczos and block conjugate gradient algorithm. Extends to symmetric indefinite problems using Gauss elimination with partial pivoting (rather than Paige and Saunders' LQ factorization). Discusses parallel implementation on SIMD machines with number of processors much less than n. Solves model problem using red/black ordering to reduce system to half size, and then applies conjugate gradients preconditioned by tridiagonal blocks. Presents similar results for nine-point operator and biharmonic. Presents a modification to the Powell (1964) nonlinear conjugate direction algorithm based on a QR factorization to prevent directions from becoming linearly dependent.

242. /NEL/ Ruhe, Axel (1974a) "Iterative Eigenvalue Algorithms for Large Symmetric Matrices," in *Numerische Behandlung von Eigenwertaufgaben Oberwolfach 1972,* ISNM 24, Birkhäuser Verlag, Basel and Stuttgart, pp. 97-115.

Compares Lanczos (1950) algorithm with optimization of the Rayleigh quotient by steepest descent and conjugate gradient algorithms on problems in which the matrix is on secondary storage units. Studies the rates of convergence and gives numerical examples showing that the Lanczos algorithm is superior, although the conjugate gradient algorithm also works well on well-conditioned problems and is easily implemented.

243. /AELN/ Ruhe, Axel (1974b) "SOR-Methods for the Eigenvalue Problem with Large Sparse Matrices," *Math. of Comp.* 28, pp. 695-710.

Applies SOR to minimization of the Raleigh quotient. Notes that it is as effective as conjugate gradient minimization if the separation of the eigenvalues is not too bad, but conjugate gradient minimization or the Lanczos (1950) algorithms are preferred in case of poor separation.

244. /AC/ Saxena, Narendra K. (1974) "Adjustment Technique without Explicit Formation of Normal Equations (Conjugate Gradient Method)," *J. of Geophysical Research* 79, pp. 1147-1152.

Applies conjugate gradients to the normal equation formulation of a geodetic triangulation system of size 965×573. References his 1972 technical report for a detailed description of programs and testing. Gives good reference list for German literature.

245. /S/ Saylor, Paul E. (1974) "Second Order Strongly Implicit Symmetric Factorization Methods for the Solution of Elliptic Difference Equations," *SIAM J. Numer. Anal.* 11, pp. 894-908.

Proposes an alternative to the SIP method of Stone which gives a symmetric (rather than unsymmetric) splitting of the matrix while preserving the "second-order" property that (discretizations of) all first degree polynomials are in the null space of the error matrix. Concludes that all such symmetric second-order splittings are impractical numerically.

246. /AC/ Sayre, D. (1974) "Least-Squares Phase Refinement. II. High-Resolution Phasing of a Small Protein," *Acta Cryst.* A30, pp. 180-184.

Uses conjugate gradients on a parametric least squares problem with only five iterations for each set of parameters but total computational costs of $7500.

247. /ANP/ Wilson, William J. (1974) "SST Flight - Profile Optimisation," *Proc. Inst. of Elect. Engrs.* 121, pp. 739-745.

Uses preconditioned conjugate gradients to solve a control problem related to the Concorde.

1975

248. /CP/ Chandra, R., S. C. Eisenstat, and M. H. Schultz (1975) "Conjugate Gradient Methods for Partial Differential Equations," in *Advances in Computer Methods for Partial Differential Equations*, ed. R. Vichnevetsky, AICA, Rutgers University, New Brunswick, New Jersey, pp. 60-64.

Applies the conjugate gradient algorithm, preconditioned by the Dupont, Kendall, Rachford (1968) scheme, to the model problem in two dimensions, and states that $O(n^{5/4}\log\varepsilon^{-1})$ operations are required to reduce the error by ε. Gives analogous results in three dimensions.

249. /CN/ Danilin, Yu. M. (1975) "Dual Direction Methods for Function Minimization," in *Optimization Techniques: IFIP Technical Conference*, ed. G. I. Marchuk, Springer-Verlag (Lecture Notes in Computer Science 27), New York, pp. 289-293.

Summarizes some previous work in developing algorithms based on biconjugate vectors. Gives a superlinearly convergent algorithm based on biorthogonal directions for minimizing functions whose Hessian matrices are uniformly positive definite with bounded condition number.

250. /EL/ Davidson, Ernest R. (1975) "The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices," *J. Comp. Phys.* 17, pp. 87-94.

Uses Lanczos (1950) to motivate a method, also related to coordinate relaxation, less expensive than Lanczos per iteration.

251. /N/ Dennemeyer, R. F. and E. H. Mookini (1975) "CGS Algorithms for Unconstrained Minimization of Functions," *J. of Optimization Theory and Applications* 16, pp. 67-85.

Uses conjugate direction algorithms with directions generated by Gram-Schmidt on a predetermined set of vectors.

252. /N/ Dixon, L. C. W. (1975) "Conjugate Gradient Algorithms: Quadratic Termination without Linear Searches," *J. Inst. Maths. Applics* 15, pp. 9-18.

Accepts the first improved point generated by a certain line search procedure rather than the function minimizer. Stores and updates two extra vectors.

253. /CN/ Gay, David M. (1975) *Brown's Method and Some Generalizations, with Applications to Minimization Problems,* TR 75-225, Ph.D. thesis, Department of Computer Science, Cornell University, Ithaca, New York.

Draws the connections among a 1966 method of Brown for solution of nonlinear equations, Craig's (1955) method for linear equations, and Stewart's (1973) generalized conjugate direction algorithms. Uses these properties to derive new algorithms for constrained and unconstrained minimization.

254. /AEL/ Hausman Jr., R. F., S. D. Bloom, and C. F. Bender (1975) "A New Technique for Describing the Electronic States of Atoms and Molecules -- The Vector Method," *Chem. Phys. Letters* 32, pp. 483-488.

Uses Lanczos (1950) to determine smallest eigenvalues of large matrices modeling configuration interaction wavefunctions.

255. /EL/ Haydock, R., V. Heine, and M. J. Kelley (1975) "Electronic Structure Based on the Local Atomic Environment for Tight-Binding Bands: II," *J. Phys. C: Solid State Physics* 8, pp. 2591-2605.

Refines their 1972 derivation of the Lanczos algorithm.

256. /C/ Hestenes, Magnus R. (1975) "Pseudoinverses and Conjugate Gradients," *Communications of the ACM* 18, pp. 40-43.

Notes that conjugate gradients can be used on least squares problems using symmetry of A^*A or AA^*. Gives an algorithm for constructing a pseudoinverse of A using the standard formulation, but with x, r, and p considered to be matrices, and with inner product $(x,y)=\sum x_{ij}y^*_{ij}$. Notes that the algorithm is $O(n^4)$ for computing the inverse of a dense nonsingular matrix.

257. /N/ Huang, H. Y. and A. K. Aggarwal (1975) "A Class of Quadratically Convergent Algorithms for Constrained Function Minimization," *J. of Optimization Theory and Applications* 16, pp. 447-485.

Derives a family of algorithms which includes the conjugate gradient algorithm and the variable metric methods. Shows that all the algorithms behave identically for quadratic minimization subject to linear constraints and terminate in at most $n-r$ steps when r is the number of linearly independent constraints.

258. /AEL/ Ibarra, R. H., M. Vallieres, and D. H. Feng (1975) "Extended Basis Shell-Model Study of Two-Neutron Transfer Reactions," *Nuclear Physics* A241, pp. 386-406.

Discusses the Lanczos method using fewer than n steps to calculate two-nucleon overlaps.

259. /CPS/ Marchuk, G. I. (1975) *Methods of Numerical Mathematics,* Springer-Verlag, New York.

(Translated by Jiri Růžička from *Metody Vychislitel'noi Mathematiki, 1973, Nauka, Novosibirsk)* Gives an exposition of the conjugate gradient method (Sec. 3.2) and suggests its use in acceleration iterative methods based on matrix splittings.

260. /CL/ Paige, C. C. and M. A. Saunders (1975) "Solution of Sparse Indefinite Systems of Linear Equations," *SIAM J. Numer. Anal.* 12, pp. 617-629.

Extends the Lanczos algorithm to nonpositive definite systems by replacing the implicit LU decomposition without pivoting with various stable factorizations. Produces a generalization of the conjugate gradient (SYMMLQ) and minimum residual (MINRES) algorithms. Notes that the algorithm can be used for linear least squares problems. Reports computational experience, and refers to a technical report for programs.

261. /AEL/ Platzman, G. W. (1975) "Normal Modes of the Atlantic and Indian Oceans," *J. of Physical Oceanography* 5, pp. 201-221.

Uses the Lanczos algorithm to find eigenvalues in a small range of the spectrum.

262. /EP/ Ruhe, Axel (1975) "Iterative Eigenvalue Algorithms Based on Convergent Splittings," *J. of Computational Phys.* 19, pp. 110-120.

Solves generalized eigenvalue problem by splitting $A-\mu B=V-H$ and iterating $x_{s+1}=V^{-1}Hx_s$.

263. /C/ Stewart, G. W. (1975) "The Convergence of the Method of Conjugate Gradients at Isolated Extreme Points of the Spectrum," *Numer. Math.* 24, pp. 85-93.

Shows that the error component in the direction of an eigenvector for an extreme and isolated eigenvalue converges rapidly.

264. /ACEL/ Todd, John (1975) "Numerical Analysis at the National Bureau of Standards," *SIAM Rev.* 17, pp. 361-370.

"This scheme was devised in 1951 simultaneously by E. Stiefel in Zürich and M. R. Hestenes at INA. There was considerable preliminary work at INA in which Forsythe, Karush, Motzkin and Rosser also participated. At the same time Lanczos, also at INA, adapted his (1950) method of minimized iterations for the determination of the characteristic polynomial of A to one for the solution of $Ax=b$ and arrived at the same conjugate direction method. There was much further work at INA by [L.] Paige, M. Stein, Hayes, Hochstrasser, L. Wilson and Curtiss. The definitive report is [Hestenes and Stiefel (1952) and Hestenes (1956)]."

265. /N/ Turner, W. C. and P. M. Ghare (1975) "Use of Dynamic Programming to Accelerate Convergence of Directional Optimization Algorithms," *J. of Optimization Theory and Applications* 16, pp. 39-47.

Gives algorithm for determining step lengths for multiple steps at once. Claims applicability to conjugate gradients but does not apply it.

266. /EL/ Underwood, Richard (1975) *An Iterative Block Lanczos Method for the Solution of Large Sparse Symmetric Eigenproblems*, Ph.D. dissertation, Stanford University Computer Science Dept. Report STAN-CS-75-496, Stanford, California.

Develops the block version of the Lanczos algorithm and generalizes Paige's convergence theory for the eigenvalue and eigenvalue estimates. Suggests strategies for choosing a blocksize. Gives Fortran implementation with full reorthogonalization.

267. /C/ Woźniakowska, G. and H. Woźniakowski (1975) "Algorytmizacja metody me-T," *Roczniki Polskiego Towarzystwa Matematycznego: Matematyka Stosowana* Seria 3,5, pp. 51-60.

"Algorithmization of the me-T method" Presents an Algol code for a combined Chebyshev and minimal error iteration applied to the normal equations for a rectangular system.

268. /C/ Woźniakowski, H. (1975) "Metoda minimalnych B-bledów dla wielkich ukladów równań liniowych o dowolnej macierzy," *Roczniki Polskiego Towarzystwa Matematycznego: Matematyka Stosowana* Seria 3,5, pp. 5-27.

"The Method of Minimal B-Errors for Large Systems of Linear Equations with an Arbitrary Matrix" develops conjugate gradient methods which minimize the B-norm of the error, and suggests applying them to normal equations.

1976

269. /ACP/ Allwright, J. C. (1976) "Conjugate Gradient Versus Contraction Mapping," *J. of Optimization Theory and Applications* 19, pp. 587-611.

Notes that preconditioned conjugate gradients using a matrix splitting requires fewer iterations than the stationary iterative method alone. Applies the conjugate gradient technique to control problems, using R or $R+\bar{W}^*Q\bar{W}$ as preconditioning for $R+W^*QW$, where $\bar{W}$ is a low rank approximation to W.

270. /CP/ Andersson, Lennart (1976) *SSOR Preconditioning of Toeplitz Matrices,* Ph.D. Thesis, Computer Sciences Department, Chalmers University of Technology, Gōteborg.

Analyzes the eigenvalue distribution for SSOR preconditioning of Toeplitz matrices and for triangular Toeplitz matrix preconditioning. Applies the results to discretized elliptic differential equations.

271. /CPS/ Axelsson, O. (1976) "A Class of Iterative Methods for Finite Element Equations," *Computer Methods in Applied Mechanics and Engineering* 9, pp. 123-137.

Discusses the preconditioned conjugate gradient algorithm in three-term recurrence form using matrix splittings such as ADI and SSOR as preconditioners, and gives operation counts for solving self-adjoint second-order problems in d dimensions by SSOR preconditioning of the conjugate gradient algorithm compared with a direct method and with standard conjugate gradients. Gives convergence bounds for preconditioned conjugate gradient algorithms when the eigenvalues fall in two disjoint intervals and when there are only a few isolated large eigenvalues.

272. /S/ Beauwens, Robert and Lena Quenon (1976) "Existence Criteria for Partial Matrix Factorizations in Iterative Methods," *SIAM J. Numer. Anal.* 13, pp. 615-643.

Presents existence criteria for partial factorizations such as Stone's method and Buleev's method. Extends results to block factorizations and studies convergence properties of iterations based on the symmetric point factorization methods.

273. /ACEL/ Cline, Alan K., Gene H. Golub, and George W. Platzman (1976) "Calculation of Normal Modes of Oceans Using a Lanczos Method," in *Sparse Matrix Computations*, ed. James R. Bunch and Donald J. Rose, Academic Press, New York, pp. 409-426.

Determines interior eigenvalues of matrix of dimension 1919 by using inverse iteration, solving the linear systems using the Lanczos decomposition or the Paige and Saunders (1975) algorithm.

274. /C/ Concus, Paul and Gene H. Golub (1976) "A Generalized Conjugate Gradient Method for Nonsymmetric Systems of Linear Equations," in *Computing Methods in Applied Sciences and Engineering*, ed. R. Glowinski and J. L. Lions, Springer-Verlag, New York, pp. 56-65.

Develops an iterative method in three-term recurrence form which requires that the symmetric part of A be positive definite and that linear systems involving it be easy to solve. Shows that estimates of the eigenvalues of $M^{-1}A$ can be obtained in the course of the iteration, and that the algorithm takes at most k iterations if there are k distinct eigenvalues. Gives a computational example.

275. /ACP/ Concus, Paul, Gene H. Golub, and Dianne P. O'Leary (1976) "A Generalized Conjugate Gradient Method for the Numerical Solution of Elliptic Partial Differential Equations," in *Sparse Matrix Computations*, ed. James R. Bunch and Donald J. Rose, Academic Press, New York, pp. 309-332.

Gives general exposition of preconditioned conjugate gradients. Suggests using conjugate gradients until approximations to extreme eigenvalues can be determined and then switching to the Chebyshev semi-iterative method. Discusses block 2-cyclic, SSOR, and sparse factorization preconditioning. Gives numerical comparison of Chebyshev and conjugate gradients on elliptic difference equations preconditioned by a discrete Laplacian and for T-shaped regions preconditioned by fast direct methods.

276. /CN/ Dennis, Jr., J. E. (1976) "A Brief Survey of Convergence Results for Quasi-Newton Methods," in *Nonlinear Programming*, ed. Richard W. Cottle and Carlton E. Lemke, American Mathematical Society, Providence, Rhode Island, pp. 185-199.

Surveys quasi-Newton algorithms and convergence results.

277. /AC/ Dodson, E. J., N. W. Isaacs, and J. S. Rollett (1976) "A Method for Fitting Satisfactory Models to Sets of Atomic Positions in Protein Structure Refinements," *Acta Cryst.* A32, pp. 311-315.

Applies conjugate gradients to a least squares problem.

278. /ACNP/ Douglas, Jr., Jim and Todd Dupont (1976) "Preconditioned Conjugate Gradient Iteration Applied to Galerkin Methods for a Mildly-Nonlinear Dirichlet Problem," in *Sparse Matrix Computations*, ed. James R. Bunch and Donald J. Rose, Academic Press, New York, pp. 333-348.

Uses a fast direct method for Poisson's equation to precondition a conjugate gradient iteration.

279. /CN/ Fletcher, R. (1976) "Conjugate Gradient Methods for Indefinite Systems," in *Numerical Analysis Dundee 1975*, ed. G. A. Watson, Springer Verlag, New York, pp. 73-89.

Studies symmetric indefinite problems. Discusses the minimum residual algorithm. Rediscovers Fridman (1963) algorithm and relates it to Paige and Saunders (1975). Discusses the Luenberger (1969) algorithm and proposes alternate formulas without fully solving the instability problem. Discusses extensions to nonlinear problems.

280. /C/ Il'in, B. P. (1976) "Some Estimates for Conjugate Gradient Methods," *USSR Comp. Math. and Math. Phys.* 16, No. 4, pp. 22-30.

(Zh. vȳchisl. Mat. mat. Fiz. 16, No. 4, pp. 847-855.) Using Lanczos polynomials instead of Chebyshev polynomials, obtains estimates of the form $(Ae_n,e_n) \leq (e_0,e_0)/(2n+1)^2$ for conjugate gradients, and analogous results for minimum residual and "minimal discrepancies" algorithms, whenever $|| A || \leq 1$. Obtains a family of bounds for the error at the kth step by using the product of the jth degree Lanczos polynomial and the $k-j$th degree Chebyshev polynomial. Applies the estimates to conjugate gradients preconditioned by ADI.

281. /EL/ Kahan, W. and B. N. Parlett (1976) "How Far Should You Go with the Lanczos Process?" in *Sparse Matrix Computations*, ed. James R. Bunch and Donald J. Rose, Academic Press, New York, pp. 131-144.

Develops error bounds for the exact algorithm and computable diagnostics for the algorithm with inexact arithmetic.

282. /AEL/ Kaplan, Theodore and L. J. Gray (1976) "Elementary Excitations in Random Substitutional Alloys," *Physical Review* B14, pp. 3462-3470.

Applies Lanczos (1950) algorithm to find eigenvalues of models of disordered systems.

283. /AC/ Konnert, John H. (1976) "A Restrained-Parameter Structure-Factor Least-Squares Refinement Procedure for Large Asymmetric Units," *Acta Cryst.* A32, pp. 614-617.

Applies conjugate gradients to a least squares problem in which some matrix coefficients are threshholded to zero.

284. /N/ Lenard, Melanie L. (1976) "Convergence Conditions for Restarted Conjugate Gradient Methods with Inaccurate Line Searches," *Math. Programming* 10, pp. 32-51.

Proves convergence of restarted conjugate gradients with inexact line searches when the second derivative matrix is continuous, bounded, and Lipschitz at the solution. Obtains n-step quadratic convergence for some conjugate gradient methods with inexact line search.

285. /L/ Paige, C. C. (1976) "Error Analysis of the Lanczos Algorithm for Tridiagonalizing a Symmetric Matrix," *J. Inst. Maths. Applics.* 18, pp. 341-349.

Gives a rounding-error analysis and relates loss of orthogonality to convergence.

286. /CN/ Powell, M. J. D. (1976a) "Some Convergence Properties of the Conjugate Gradient Method," *Math. Programming* 11, pp. 42-49.

Proves that conjugate gradients on a quadratic objective function with arbitrary downhill initial direction is either finitely terminating or linearly convergent.

287. /CN/ Powell, M. J. D. (1976b) "Some Global Convergence Properties of a Variable Metric Algorithm for Minimization without Exact Line Searches," in *Nonlinear Programming*, ed. Richard W. Cottle and Carlton E. Lemke, American Mathematical Society, Providence, Rhode Island, pp. 53-72.

Studies the convergence of the BFGS algorithm without exact line searches. Shows convergence for convex functions and superlinear convergence if the second derivative matrix is positive definite at the solution under some conditions on the search.

288. /AC/ Proskurowski, Wlodzimierz and Olof Widlund (1976) "On the Numerical Solution of Helmholtz's Equation by the Capacitance Matrix Method," *Math. of Comp.* 30, pp. 433-468.

Uses the SYMMLQ algorithm of Paige and Saunders (1975) as an iterative method on a dense well-conditioned matrix. Attributes the success of a small number of iterations to a special eigenvalue distribution. Notes that the matrix can be represented in product form, as in George (1970).

289. /CN/ Safro, V. M. (1976) "The Rate of Convergence of Some Gradient Methods," *USSR Comp. Math. and Math. Phys.* 16, No. 2, pp. 212-215.

(Zh. vy̌chisl. Mat. mat. Fiz. 16, No. 2, pp. 496-499.) Obtains convergence estimates for a Fletcher-Reeves-type conjugate gradient algorithm under assumptions that the differentiable function to be minimized in Hilbert space is convex or strongly convex.

290. /AC/ Squire, Wm. (1976) "The Solution of Ill-Conditioned Linear Systems Arising from Fredholm Equations of the First Kind by Steepest Descents and Conjugate Gradients," *Internat. J. for Numerical Methods in Engineering* 10, pp. 607-617.

Gets smooth solutions to a dense system by taking several conjugate gradient steps.

291. /AL/ von Charzewski, C. and R. M. Dreizler (1976) "Constrained Hartree-Fock and Minimization of the Variance," *Z. Physik A* 278, pp. 35-40.

Compares the performance of calculations minimizing both the unprojected and the projected expectation value and variance of the Hamiltonian, both calculations involving the Lanczos (1950) algorithm.

Author Index