(3)

Do We Fully Understand the Symmetric Lanczos Algorithm Yet? *

Beresford N. Parlett †

Abstract

Imagine that one has computed the real n-vectors $b, Ab, A^2b, \ldots, A^{m-1}b$ where A is a real symmetric $n \times n$ matrix. Lanczos showed us in 1950 how to construct a much better basis for the (Krylov) space spanned by these power vectors and for little extra cost. The new basis $\{q_1, q_2, \ldots, q_m\}$, now called the Lanczos basis, has two nice properties: (i) it is orthonormal, (ii) the representation of A's projection is a symmetric tridiagonal matrix T_m. Property (ii) is synonymous with the three term recurrence governing the Lanczos vectors. Moreover some of T_m's eigenvalues, called Ritz values hereafter, are excellent approximations to some of A's eigenvalues even when $m << n$. In addition we can tell, with little expense, which Ritz values are also eigenvalues. One surprising implication of these properties is that it is easier to find the largest few eigenvalues of A than to solve $Ax = b$!

When the Lanczos algorithm is implemented in a computer the user discovers an unpleasant fact. Property (i) fails completely for m as small as 20 or 30 and consequently the computed T_m's relation to A is unclear. Lanczos was aware of this blemish and proposed the obvious remedy: keep applying the Gram-Schmidt process to each new Lanczos vector as it is computed. The catch here is that all the $\{q_i\}$ must be kept handy whereas in exact arithmetic only the three latest Lanczos vectors are needed and earlier q's may be discarded. The arithmetic cost of this full reorthogonalization grows quadratically with m. So the hope of computing T_n efficiently and accurately by the Lanczos algorithm was dashed and other methods prevailed. In exact arithmetic T_n is similar to A and the algorithm stops.

What Lanczos and Wilkinson both failed to see is that there is structure in the way that orthogonality is lost. This structure is revealed by a clever change of basis and it was discovered by C. C. Paige in 1969/1970 while writing his dissertation. Moreover the computed T_m retains information about A. Thus the loss of orthogonality delays the discovery of A's eigenvalues by the simple Lanczos algorithm but does not prevent the attainment of full accuracy if enough steps are taken. In finite precision arithmetic the simple Lanczos algorithm will run forever and we are just beginning to come up with good models that describe how T_m relates to A when $m >> n$.

1 Introduction

Today the Lanczos algorithm seems so natural, so inevitable, and so simple that it is difficult to imagine that it was not part of the numerical scene until 1950. Of course digital computers did not appear until the end of World War II but, equally important, is the fact that the concept of a tridiagonal matrix was not in the mental tool box of a scientist in 1950. A tridiagonal matrix does not reveal its eigenvalues or eigenvectors to the casual eye

*Supported by ONR, contract N000014-90-J-1372.

†Department of Mathematics, University of California, Berkeley, CA 94720, USA.

so, at best, it seemed a mere stepping stone to the desired spectral information. Indeed it was not until 1954 that safe and efficient ways of computing the spectrum of a symmetric tridiagonal matrix were discovered. For some applications, such as solving large systems of differential equations, a tridiagonal representation may suffice and diagonalization is overkill, see [13], [8].

This essay will discuss the Lanczos algorithm in the context of exact arithmetic and in the context of computer arithmetic. A sequence of pictures is presented that reveals, in striking detail, the structure governing the finite precision algorithm.

The essay ends with a brief discussion of recent work designed to bound the number of steps needed to determine a given 'well represented' eigenvalue to working accuracy. Indeed it is not at all obvious that the tridiagonal matrix produced by the Lanczos algorithm will eventually have a spectrum that comes arbitrarily close to each eigenvalue of the given matrix, even if the algorithm is run forever.

Let us now introduce the notation that is needed to tell our story. We start with an $n \times n$ Hermitian matrix A produced by a physicist or engineer. Either all or part of A's spectrum is wanted and the challenge is that n may be large; $n = 10^4$ is common, $n = 10^6$ no longer makes headlines, and $n = 10^8$ is waiting around the corner. In several ways it helps to think of A as a self-adjoint linear operator and to forget n.

Physicists are in the habit of talking glibly about diagonalizing an Hermitian matrix because their instructors tell them that this is always possible. The diagonal matrix is called Λ and we write

$$A = Z\Lambda Z^* \tag{1}$$

where Z is a unitary matrix whose columns are eigenvectors of A. For any complex object C we write C^* for its conjugate transpose. In this essay we are concerned only with the first step in the task of computing Λ, namely the production of a real tridiagonal matrix T where

$$T = T_n = \begin{bmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \beta_3 & \cdot & \cdot & \\ & & \cdot & \cdot & \beta_n \\ & & & \beta_n & \alpha_n \end{bmatrix}, \quad \beta_i > 0, \quad i = 2, \ldots, n$$

$$= tridiag\left(\begin{matrix} & \beta_2 & & \beta_3 & & \cdot & & \beta_n & \\ \alpha_1 & & \alpha_2 & & \cdot & & \cdot & & \alpha_n \\ & \beta_2 & & \beta_3 & & \cdot & & \beta_n & \end{matrix} \right) \tag{2}$$

so that

$$A = QTQ^* \tag{3}$$

with $Q^* = Q^{-1}$ being unitary.

Is T canonical in the same way as Λ? Certainly not. There are infinitely many essentially different T matrices, a family with $n-1$ degrees of freedom. This freedom can be specified very nicely.

THEOREM 1.1. *If Equation (3) above holds and Q is written by columns as $Q = [q_1, \ldots, q_n]$ then both T and Q are completely determined by A and q_1 (or by A and q_n).*

The proof of this theorem yields the Lanczos algorithm! The proof supposes that all arithmetic operations are performed exactly and numbers represented exactly. Before presenting the proof we hasten to add that this is not the way Lanczos himself introduced his algorithm. For that story see the commentaries on the 1950 paper.

Proof. To prove the theorem it is only necessary to equate columns on each side of the equation

$$AQ = QT.$$

Column 1 gives $Aq_1 = q_1\alpha_1 + q_2\beta_2$. Orthogonality of Q's columns yields $q_1^*Aq_1 = q_1^*q_1\alpha_1$ and the fact that each column of Q has Euclidean length 1 gives $\alpha_1 = q_1^*Aq_1$ and $\beta_2 = \|Aq_1 - q_1\alpha_1\|$. Throughout this essay $\|v\| := \sqrt{v^*v}$. Thus $q_2 = (Aq_1 - q_1\alpha_1)/\beta_2$.

In the same way, on equating column k on each side of $AQ = QT$ one finds that

$$\begin{aligned}\alpha_k &= q_k^*Aq_k = q_k^*(Aq_k - q_{k-1}\beta_k),\\ \beta_{k+1} &= \|Aq_k - q_{k-1}\beta_k - q_k\alpha_k\|,\\ q_{k+1} &= (Aq_k - q_{k-1}\beta_k - q_k\alpha_k)/\beta_{k+1}.\end{aligned}$$

So if q_{k-1}, β_k, and q_k are known then α_k, β_{k+1}, and q_{k+1} are completely determined. Notice that we rather cleverly *assumed* that all β values were positive. In practice, if $\beta_{k+1} = 0$, we are delighted because then the linear span of $\{q_1, q_2, \ldots, q_k\}$ is invariant under A and every eigenvalue of T_k is an eigenvalue of A.

Could anything be more simple?

Let us mention another attractive feature of the algorithm. We live in an age that measures standard of living by the amount that the average citizen discards. We love to throw things away. Notice that at the end of step k we may throw away q_{k-1}; it has served its purpose. In order to compute eigenvalues it is only necessary to store T and to use 3 n-vectors in the fast memory. The arithmetic effort at each step is dominated by the formation of Aq_k. Indeed, if A is a sparse matrix and the cost of forming Aq_k is $O(n)$ rather than $O(n^2)$ then we have an algorithm that computes T_n with $O(n^2)$ effort.

It is too good to be true–as we shall soon see.

2 Rayleigh-Ritz Approximations

There is more to be said about the Lanczos algorithm in exact arithmetic. Two matrix equations specify the output of the algorithm *at each step*. Define

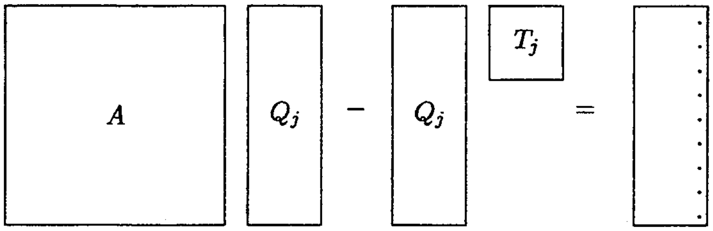
$$\begin{aligned}Q_j &:= [q_1, q_2, \ldots, q_j],\\ e_j^* &:= (0, 0, \ldots, 0, 1), \quad \text{a row } j-\text{vector.}\end{aligned}$$

Then

$$(4) \qquad Q_j^*Q_j = I_j \qquad \text{(orthogonality)},$$

$$(5) \qquad AQ_j - Q_jT_j = q_{j+1}\beta_{j+1}e_j^* \qquad \text{(3-term recurrence)}.$$

Here is a picture of (5)



Up to this point we have neglected an important property of the algorithm. After an initial phase is over, i.e. for large enough j, some eigenvalues of T_j are good approximations to some eigenvalues of A. Which are the good approximations and how good are they?

There is a beautiful answer to these questions.

The extra notation needed at this point is spectral factorization, or diagonalization, of T_j:

$$T_j S_j = S_j \Theta_j \tag{6}$$

where

$$S_j = [s_1^{(j)}, s_2^{(j)}, \ldots, s_j^{(j)}], \quad S_j^{-1} = S_j^* \quad \text{is a real orthogonal } j \times j \text{ matrix,}$$

$s_i^{(j)}(k)$ denotes the kth entry in the normalized eigenvector of $\theta_i^{(j)}$. Also

$$\Theta_j = diag(\theta_1^{(j)}, \ldots, \theta_j^{(j)}),$$

where

$$\theta_1^{(j)} < \theta_2^{(j)} < \ldots < \theta_j^{(j)}$$

are the eigenvalues of T_j which, henceforth, will be called *Ritz values* (of A at step j). Associated with $\theta_i^{(j)}$ is its *Ritz vector*

$$y_i^{(j)} = Q_j s_i^{(j)} = \sum_{k=1}^{j} q_k s_i^{(j)}(k), \quad i = 1, \ldots, j. \tag{7}$$

The point of all this is that the pairs $(\theta_i^{(j)}, y_i^{(j)})$, $i = 1, \ldots, j$ are the Rayleigh-Ritz approximations to A's eigenpairs from the range of Q_j. It is known that, in several ways, these are collectively the best set of approximations from

$$range\ Q_j = span(q_1, Aq_1, A^2 q_1, \ldots, A^{j-1} q_1).$$

For more details on this subject see [1] and [11].

To distinguish the good Ritz values from the bad ones it is only necessary to postmultiply (5), the matrix form of the 3 term recurrence, by $s_i^{(j)}$ to obtain

$$AQ_j s_i^{(j)} - Q_j T_j s_i^{(j)} = q_{j+1} \beta_{j+1} s_i^{(j)}(j), \quad i = 1, \ldots, j. \tag{8}$$

By (6), $T_j s_i^{(j)} = s_i^{(j)} \theta_i^{(j)}$, and by (7)

$$Ay_i^{(j)} - y_i^{(j)} \theta_i^{(j)} = q_{j+1} \beta_{j+1} s_i^{(j)}(j) \tag{9}$$

and, on taking norms,

$$\|Ay_i^{(j)} - y_i^{(j)} \theta_i^{(j)}\| = \beta_i^{(j)} := \beta_{j+1} |s_i^{(j)}(j)|. \tag{10}$$

This is a remarkable and rare situation; one can determine the ith residual norm, i.e. $\|Ay_i^{(j)} - y_i^{(j)} \theta_i^{(j)}\|$ without forming $y_i^{(j)}$. That is good news because $y_i^{(j)}$ will not be easy to compute if all the Lanczos vectors have been discarded except for the last three. The right side of (10) requires an algorithm to update Ritz values and last entries in certain eigenvectors of a $j \times j$ tridiagonal matrix. When $j = 40$ and $n = 10^4$ the difference in arithmetic effort is significant.

Well known eigenvalue bounds, see [11] and [5], guarantee that there is at least one eigenvalue λ of A that satisfies

$$|\lambda - \theta_i^{(j)}| \leq \beta_i^{(j)}$$

and, in fact,

$$\begin{aligned} |\lambda - \theta_i^{(j)}| &\leq \beta_j^{(j)^2}/gap(i), \\ gap(i) &= \min_{\mu} |\mu - \theta_i^{(j)}|, \end{aligned}$$

where μ is an eigenvalue of A distinct from λ.

What this means in practice is that it is the quantities $|s_i^{(j)}(j)|$, $i = 1, \ldots, j$ which measure the quality of $\theta_i^{(j)}$. One can judge $y_i^{(j)}$ before computing it. There are two ways to compute $y_i^{(j)}$. Either save the old q-vectors $q_1, \ldots, q_{j-2}$ on secondary storage or run the Lanczos algorithm for a second time accumulating $y_i^{(j)}$ alongside q_k, $k = 1, 2, \ldots, j$.

We now have two orthonormal bases for the Krylov subspace *span* $(q_1, Aq_1, \ldots, A^{j-1}q_1)$; the Lanczos basis $\{q_i\}$ and the Ritz basis $\{y_i^{(j)}\}$. Note that, in principle, the *whole* Ritz basis changes at each step. The way in which the $\theta_i^{(j)}$ approach A's spectrum has received considerable attention, see [11] and [12]. The larger the ratio $\min\{\lambda_{i+1} - \lambda_i,\ \lambda_i - \lambda_{i-1}\}/(\lambda_n - \lambda_1)$ the more rapidly does a Ritz value settle on to λ_i. Even for a randomly chosen q_1 it is not uncommon for several extreme eigenvalues to be approximated to 8 correct decimals after 30 or 40 steps, independent of n.

As we said before: It all seems too good to be true.

3 Finite Precision Arithmetic

Digital computers discard information at almost every arithmetic operation. When only the leading 50 or 60 bits of each floating point number are retained then the beautiful relationships expounded in Sections 1 and 2 fail. In order to describe what happens we change notation slightly and let Q_j and T_j denote the quantities actually stored in the computer. In contrast S_j and Θ_j denote the exact spectral factors of the computed T_j.

The orthogonality equation (4) is replaced by

$$Q_j^* Q_j = C_j^* + I_j + C_j \tag{11}$$

where C_j is strictly upper triangular. Since the computed q_i are not exactly normalized I_j should be replaced by some diagonal Δ_j that is exceedingly close to I_j but such veracity is not cost effective. The three term recurrence (5) becomes

$$AQ_j - Q_j T_j = q_{j+1}\beta_{j+1}e_j^* + F_j \tag{12}$$

where f_i, column i of F_j, is just the amount by which the 3 term recurrence fails to hold for computed q_i. It turns out that F_j remains at the round off level whereas $\|C_j\|$ grows rapidly towards 1 as j increases. We can think of (12) as (5) contaminated with 'white noise' that remains at the round off level, i.e. $\|f_i\|$ is proportional to round off unit and is independent of i.

This perturbation to (5) has significant consequences. The algorithm driven by (12) never terminates. When $j > n$ the columns Q_j cannot be linearly independent, let alone orthonormal. In fact linear independence is lost long before j is close to n. The next question is: how does T_{2n} relate to A? In exact arithmetic T_n is similar to A but that is impossible for T_{2n}.

The best way to understand this orthogonality loss is through pictures. The figures that follow show C_j as a function of two variables; one plots $|C(row,\ col)|$ as a function of $(row,\ col)$.

Figure 1 shows a typical case; orthogonality leaks away. Close to the diagonal C is negligible but at step 30 $C(k,\ k+15) = O(1)$ for most values of $k \leq 15$.

One cannot tell from Figure 1 whether linear independence has been lost yet but orthogonality loss is total.

It is not disrespectful to say that Lanczos himself, and J. H. Wilkinson, the leading expert in matrix computations from 1960-1984, both panicked at this phenomenon. Each of them insisted on doing what we now call full reorthogonalization. With this modification at step j q_{j+1} is explicitly orthonormalized against *all* preceding Lanczos vectors. See [7], p. 271, and [14], Section 38, line 3. This precaution increases the storage requirements and the arithmetic effort. No longer is the cost of *step j* constant, now it grows linearly with j. Thus full reorthogonalization seems to constrain users (except for rich physicists) to making short Lanczos runs whereas Krylov subspace theory reveals the approximating power of long Lanczos runs.

The technique of restarting is not a bad response to the difficulties but neither is it fully satisfactory. The number of extra applications of the operator beyond those needed by one long Lanczos run can be significant and is not reported by those developing restart methods.

4 Hidden Structure

Despite appearances the loss of orthogonality among the Lanczos vectors is far from random. A strong idea as to what is happening is given by changing bases from Lanczos to Ritz. The reader is invited to contemplate carefully Figures 2 through 5 which show steps 18–22 in a typical Lanczos run. The top half shows C_j and the bottom half shows the strictly upper triangular part of $Y_j^* Y_j$, the Ritz picture. Step 18 is quite revealing. Orthogonality, judged by human eyesight, has been maintained beautifully among the Ritz vectors EXCEPT that $y_{18}^{(18)}$ is a copy of $y_{17}^{(18)}$, very nearly. That is not obvious from the Lanczos picture just above. Thus range Q_{18} had dimension 17.

Step 19 (Fig. 3) seems to tell the same story, except that now $y_{19}^{(19)}$ is a copy of $y_{18}^{(19)}$ Here it is useful to remember that $y_j^{(j)}$ is the Ritz vector for the largest Ritz value;

$$\theta_1^{(j)} < \theta_2^{(j)} < \ldots < \theta_j^{(j)}.$$

Thus each of $y_{18}^{(18)}$ and $y_{19}^{(19)}$ is very close to the dominant eigenvector of A.

Step 20 (Fig. 4) spoils the simplicity of Step 19 but Step 21 restores it (almost) but space limitations forced us to omit the picture. Step 22 (Fig. 5) is very like Step 21; there are 20 orthogonal Ritz vectors plus a spare copy of the two Ritz vectors that have converged to the two dominant eigenvectors.

The idea of changing bases originated with C.C. Paige and was described in his Ph.D. dissertation in 1970/71 (London University) and in [9] and [10]. Actually neither the Lanczos basis alone nor the Ritz basis alone tells the story in its simplest terms. The key idea is to look at the angle between q_{j+1} and the previous Ritz vectors $(y_1^{(j)}, \ldots, y_j^{(j)})$. This information is available at the end of step j.

The next set of pictures, Figures 6 to 10 reveal Paige's discovery. We repeat the previous run; the picture plots $|y_i^{(j)*}q_{j+1}|$ as a function of $(i,\, j+1)$. It is helpful to remember that q_{j+1} is being tested against a different basis than is q_j.

When this is understood one notices that, in Fig. 6, q_{18} is bent towards $y_{17}^{(17)}$ and, to a smaller extent q_{17} is bent towards $y_{16}^{(16)}$. Yet both these Ritz vectors are good approximations to A's dominant eigenvector. Thus q_{17} and q_{18} are pulled in the *same* direction, q_{18} more q_{17}. Step 19 (Fig. 7) is a little disconcerting, at first, because q_{19} has maintained better orthogonality than did q_{18}. It is true; there is nothing monotonic in this orthogonality loss. By Step 22 we can see that $\theta_{20}^{(22)}$ is on its way to becoming a third copy of the dominant eigenvector!

There is a beautiful result behind the foregoing discussion.

THEOREM 4.1 (PAIGE'S THEOREM). *If local orthogonality is maintained in the Lanczos process, governed by (12), i.e.*

if $q_{k+1}^* q_k = 0$ *for all* k, *then for each* $i \le j$,

(a) $y_i^* q_{j+1} = \dfrac{\gamma_{ij}}{\beta_{j+1} s_i(j)} = \dfrac{\gamma_{ij}}{\beta_i^{(j)}}$,

(b) for $k \le j, \; i \neq k$,

$$(\theta_i - \theta_k) y_i^* y_k = \gamma_{ii}\left(\frac{s_k(j)}{s_i(j)}\right) - \gamma_{kk}\left(\frac{s_i(j)}{s_k(j)}\right) + (\gamma_{ki} - \gamma_{ik}).$$

where

$$\Gamma^{(j)} = (\gamma_{ik}) = S_j^* \,[\textit{upper triangle } (Q_j^* F_j - F_j^* Q_j)] S_j$$

and upper triangle (M) *denotes the strictly upper triangular part of* M *with the rest of the matrix filled with zero entries.*

For proofs see [9], [10] and [11].

The striking feature of part (a) is that the denominator is precisely the formula for $\|Ay_i - y_i\theta_i\|$ in exact arithmetic.

5 The Lanczos Phenomenon

Paige's work stimulated a variety of implementations of the Lanczos algorithm which differ in the extent to which orthogonality of the $\{q_i\}$ is maintained. However that is not the subject of this essay.

Experience with the simple, minimal effort, algorithm revealed that Ritz values cluster very closely round the eigenvalues of A as the number of steps increases. There will be perhaps hundreds of Ritz values within $O(\epsilon\|A\|)$ of the first eigenvalue to be found before the last one has a single Ritz value beside it. Nevertheless all eigenvalues are found eventually, thanks to round off, even when the starting vector is orthogonal to an eigenvector of A. The only known exception is the artificial case when A is diagonal and the starting vector has a zero entry.

At any particular value of j there may be Ritz values that are not eigenvalues of A but the corresponding $s_i^{(j)}(j)$ values will not be small and, sooner or later, that Ritz value $\theta_i^{(j)}$ will move. The way in which A's spectrum is revealed by Ritz value behavior is a complicated function of the starting vector, its relation to A's eigenvectors, and the distribution of A's spectrum.

Anne Greenbaum, in [3] produced a backward error analysis which showed that for a given number of steps J and a computed tridiagonal T_J there is a matrix A' and a starting vector q_1' such that the exact Lanczos algorithm applied to A', q_1' will produce T_J. This work, and further extensions of it, are described in this volume in the section written by Greenbaum.

In [2] came the first proofs of the Lanczos phenomenon: all eigenvalues are captured eventually. The results are full of technical details and we state them informally here.

THEOREM 5.1 (DRUSKIN AND KNIZHERMAN). *Let* $Az_r = z_r\lambda_r, \quad \|z_r\| = 1$. *If* $\|A\| \leq 0.9$, *if the round off unit is small enough compared to* j, *if* $\phi_r := q_1^* z_r \neq 0$, *and if* $\ln(17j/\phi_r^2) < j - 2$, *then for some* $i \leq j$

$$|\theta_i^{(j)} - \lambda_r| < \frac{\ln(17j/\phi_r^2)}{2j} = O\left(\frac{\ln j}{j}\right), \quad j \longrightarrow \infty$$

The important point here is that, in exact arithmetic, the analogous error bound is $O(1/j)$ for $j < n$. The results reflect worst case situations.

A much more realistic bound is derived in terms of the gap γ_r separating λ_r from the rest of the spectrum. The bound is too complicated to give insight of itself but it may be used to generate useful diagrams. It seems pointless to copy the result in detail. It says that under a lot of apparently reasonable conditions on ϵ there is an index i such that

$$|\theta_i^{(j)} - \lambda_r| < \kappa$$

for any κ that satisfies an inequality of the form

$$\begin{aligned} \frac{\phi_r^2\kappa}{j} \ \geq \ & |\phi_r|\sqrt{j}\epsilon_1 + 2(1 + \sqrt{j}\epsilon_1) \cdot \\ & \left\{\sqrt{\frac{2j\epsilon_3}{\kappa}} + 2\sqrt{\pi}j^2\epsilon_1 + (1 + \sqrt{j})/T_{j/2}\left(\frac{1+\gamma_r^2}{1-\gamma_r^2}\right)\right\} \end{aligned}$$

where T_k denotes the Chebyshev polynomial of the first kind of degree j. Here ϵ_1 and ϵ_3 are functions of ϵ. Note the presence of κ on both sides of the inequality. For a small enough round off unit the bound reduces to

$$|\theta_i^{(j)} - \lambda_r| < \frac{2j(1+\sqrt{j})}{\phi_r^2 T_{j/2}\left(\frac{1+\gamma_r^2}{1-\gamma_r^2}\right)}$$

which reminds specialists of the Kaniel-Paige-Saad bounds in exact arithmetic, see [12].

Knizherman [6] has further recent results (private communication) which show that the Ritz values cluster in intervals of radius $O(\epsilon\|A\|)$ about the eigenvalues. This is a significant improvement on earlier results of the form $O(\sqrt{\epsilon}\|A\|)$.

References

[1] J. K. Cullum and R. A. Willoughby , *Lanczos Algorithms for Large Symmetric Eigenvalue Computations, vol. I. Theory*, Birkhäuser, Boston, 1985.

[2] V. L. Druskin and L. A. Knizherman, *Error bounds in the simple Lanczos procedure for computing functions of symmetric matrices and eigenvalues*, Comput. Maths. Math. Phys., 31 (1991), pp. 20–30.

[3] A. Greenbaum, *Behavior of slightly perturbed Lanczos and conjugate-gradient recurrences*, Linear Algebra Appl., 113 (1989), pp. 7–63.

[4] A. Greenbaum and Z. Strakos, *Predicting the behavior of finite precision Lanczos and conjugate gradient computations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 121–137.

[5] R. A. Horn and Ch. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1988.

[6] L. A. Knizherman, *Quality of approximations to a well separated eigenvalue and location of Ritz values in the simple Lanczos process*, preprint, 1993.

[7] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Nat. Bur. Stand., 45 (1950), pp. 255-282.

[8] B. Nour-Omid and R. W. Clough, *Dynamic analysis of structures using Lanczos coordinates*, Earthquake Engineering and Structural Dynamics, 12 (1984), pp. 565–577.

[9] C. Paige, *Error analysis of the Lanczos algorithms for tridiagonalizing a symmetric matrix*, J. Inst. Math. Applc., 18 (1976), pp. 341-349.

[10] ——, *Accuracy and effectiveness of the Lanczos algorithm for the symmetric eigenproblem*, Linear Algebra and Its Applc., 34 (1980), pp. 235-258.

[11] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice Hall, Englewood Cliffs, NJ, 1980.

[12] Y. Saad, *On the rates of convergence of the Lanczos and block Lanczos methods*, SIAM J. Num. Anal., 17 (1980), pp. 687–706.

[13] E. L. Wachspress, *A numerical technique for solving group diffusion equations*, Nuclear Science and Engineering, 8 (1960), pp. 164–170.

[14] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

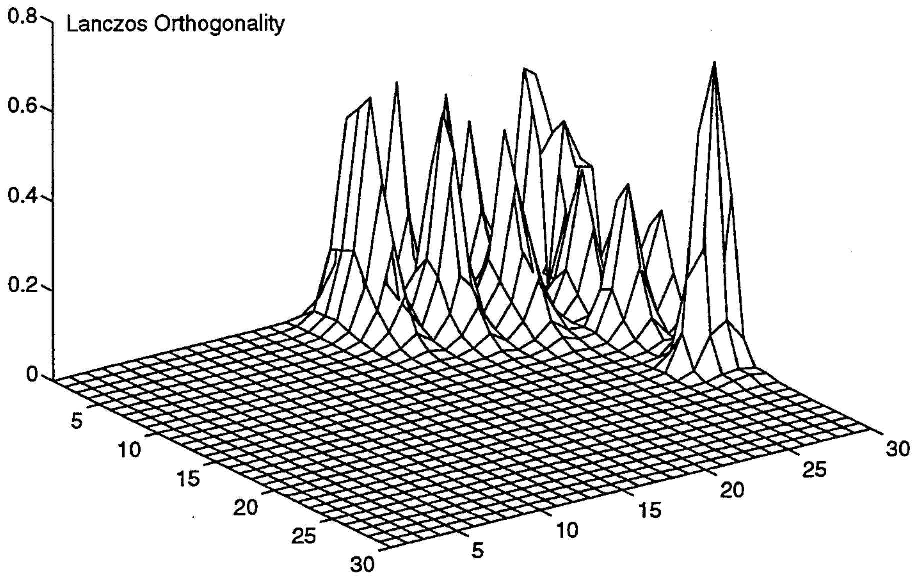


FIG. 1. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 30*

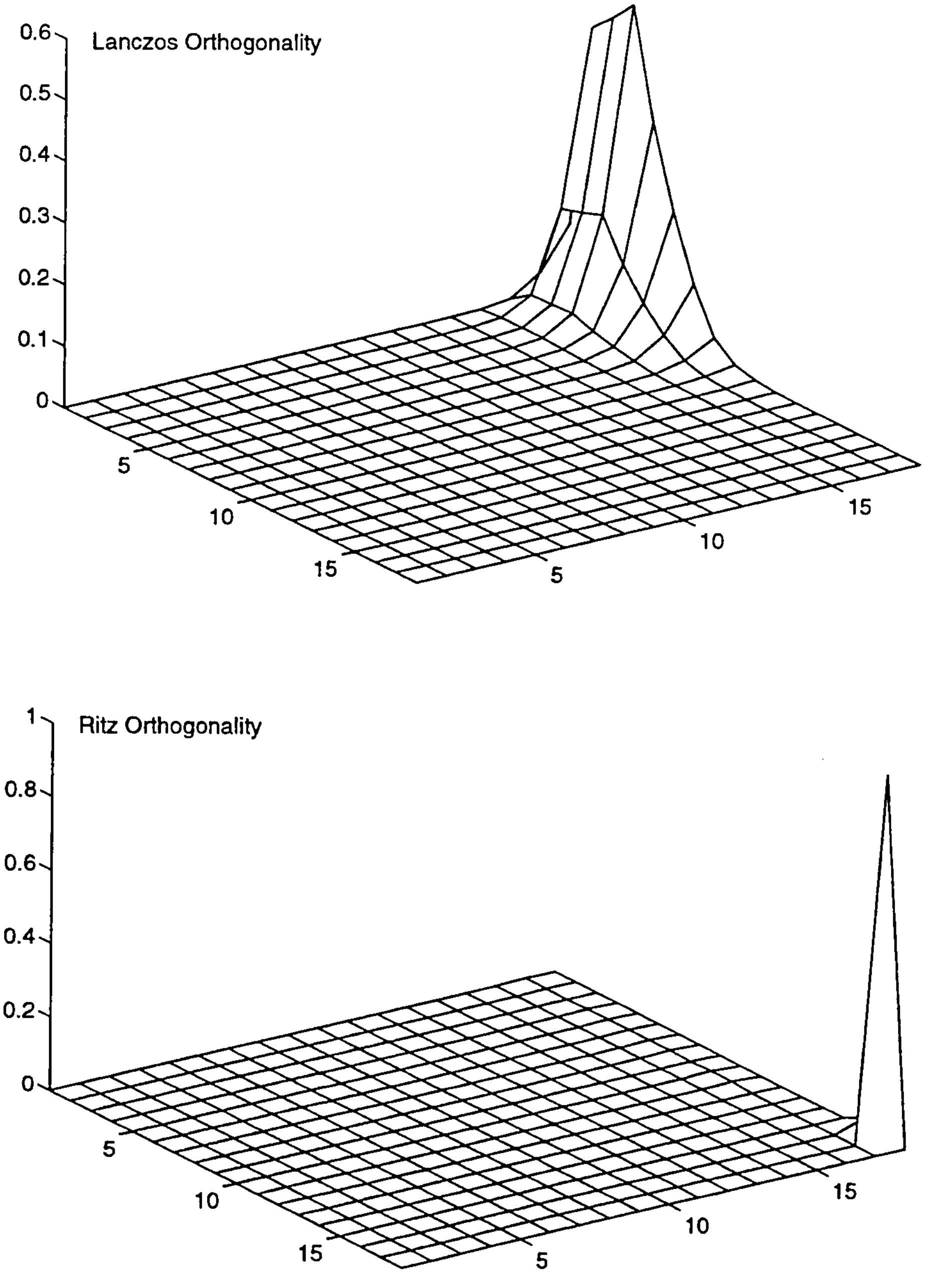


FIG. 2. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 18*

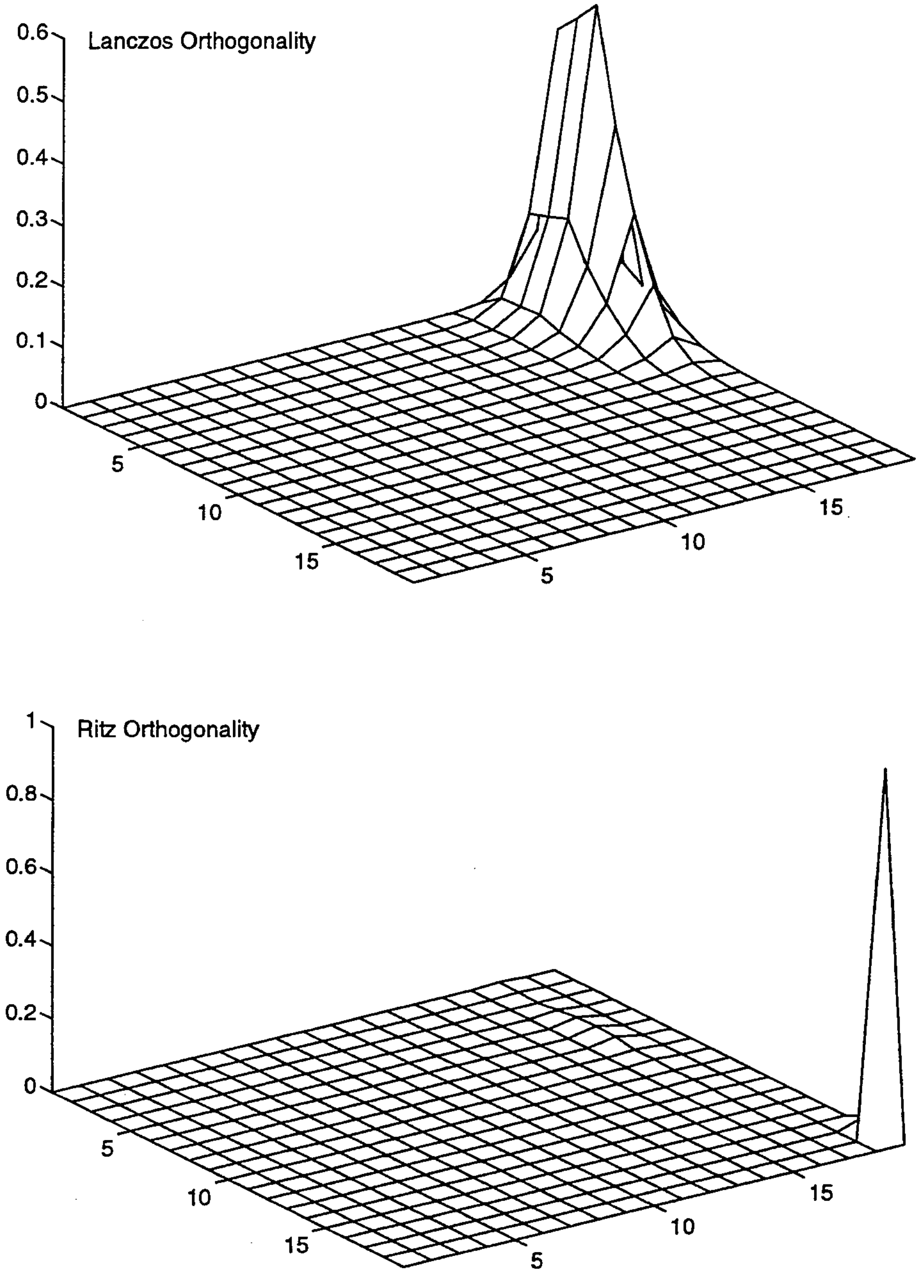


FIG. 3. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 19*

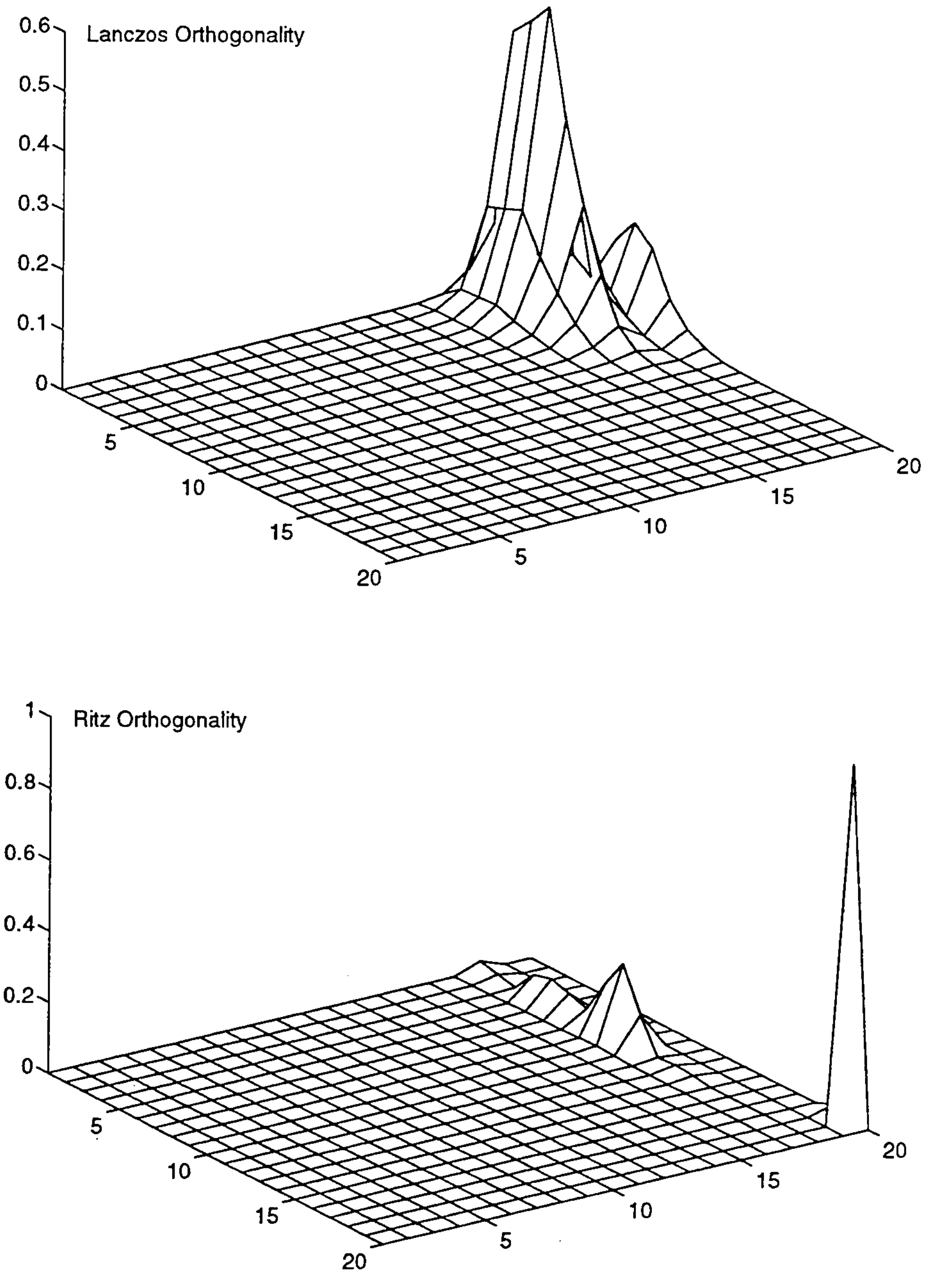


FIG. 4. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 20*

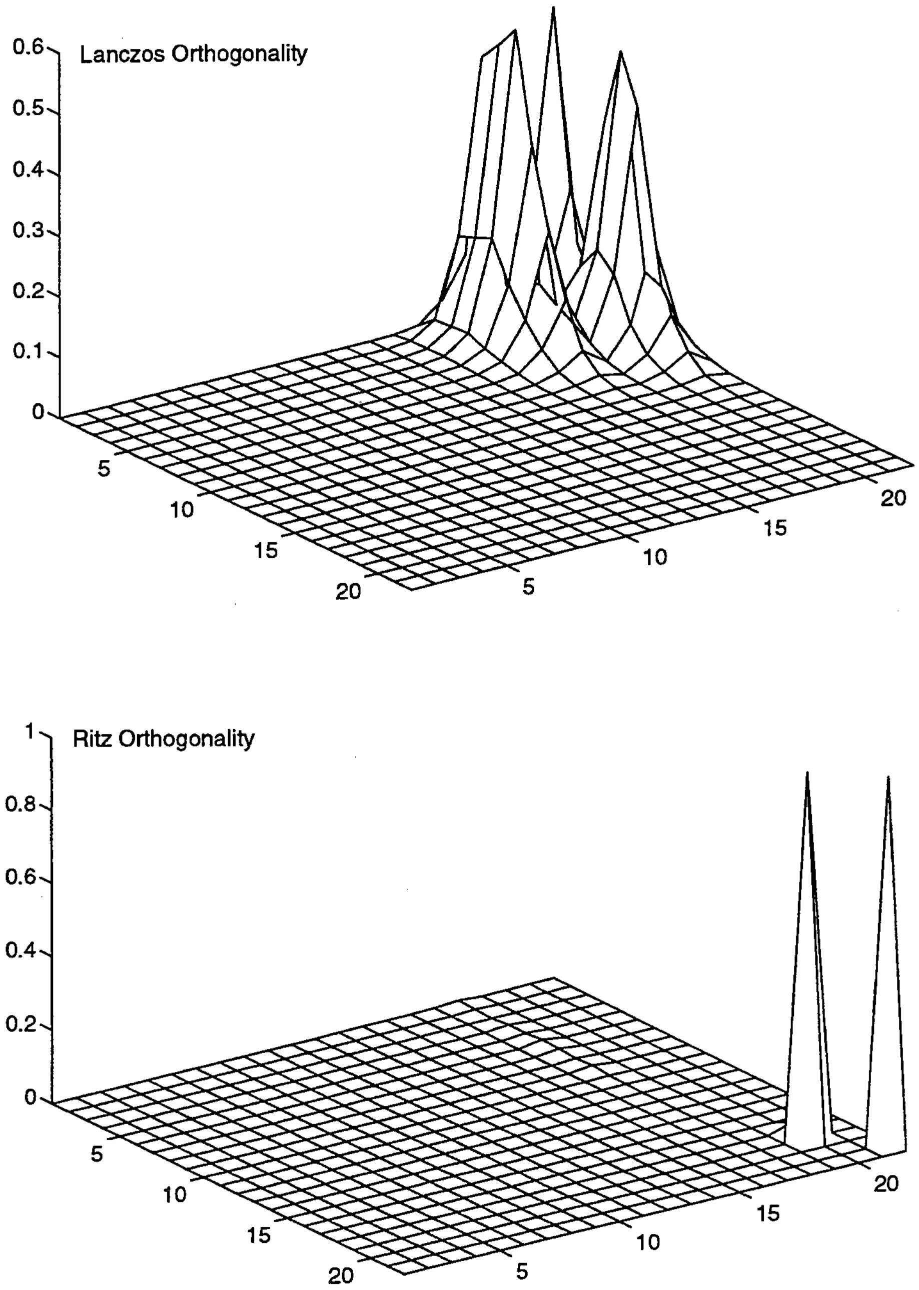


FIG. 5. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 22*

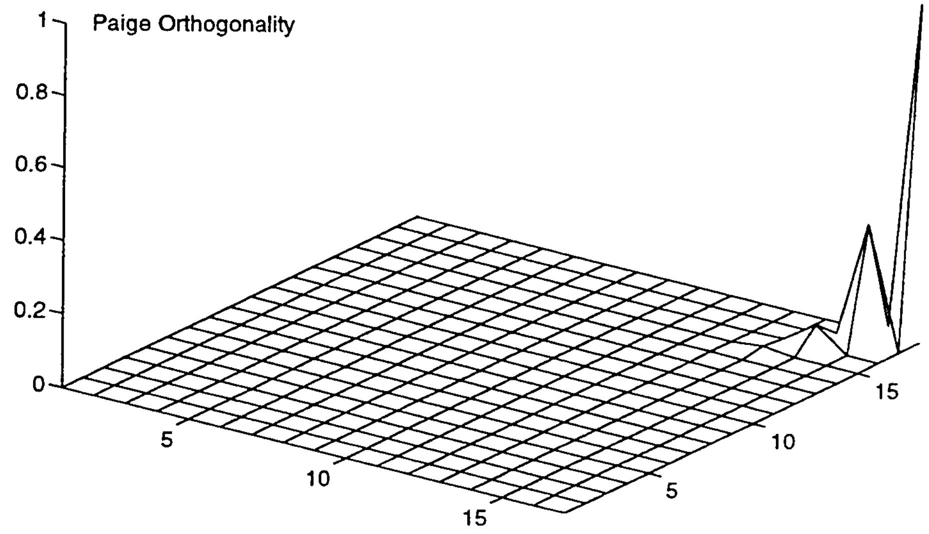


FIG. 6. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 18*

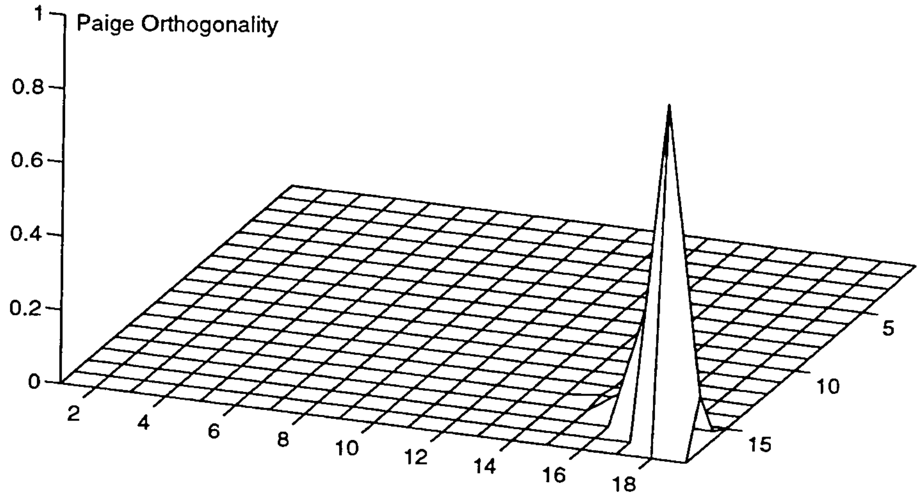


FIG. 7. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 19*

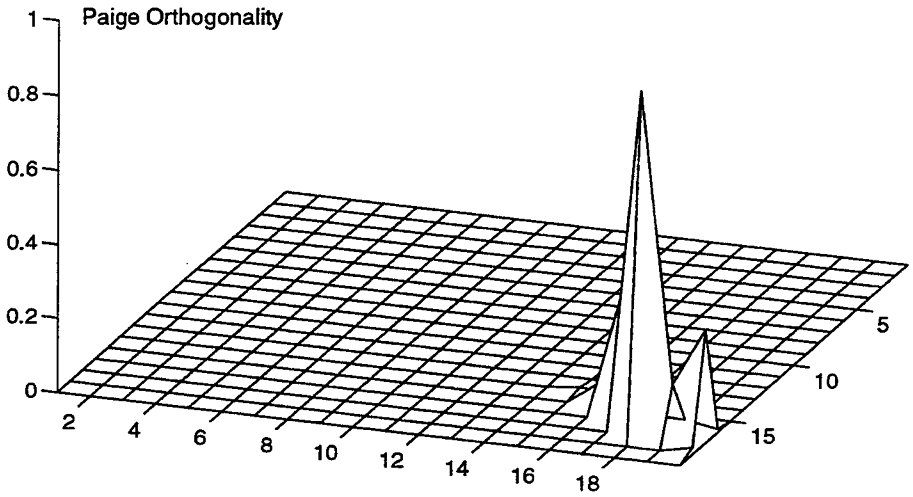


FIG. 8. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 20*

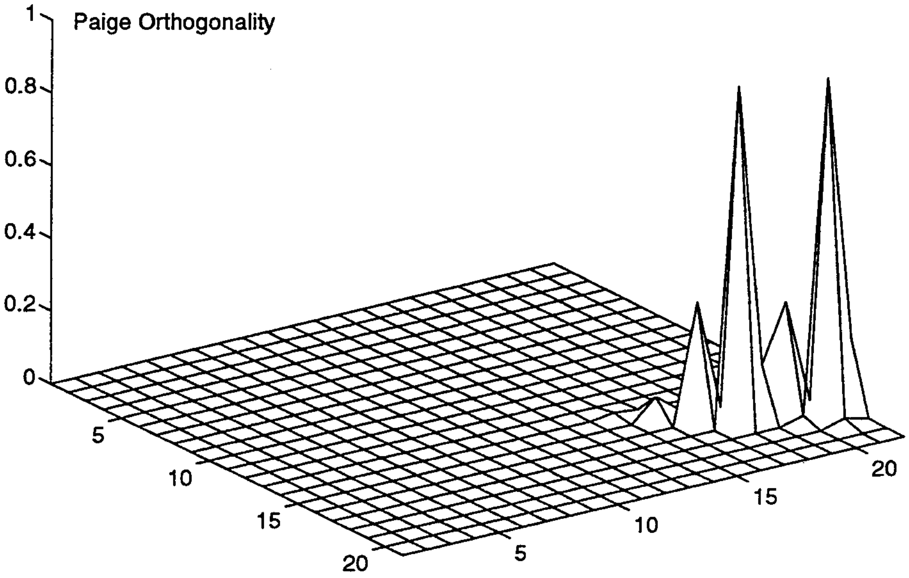


FIG. 9. *n = 100 ratio = 0.87 rel gap = 0.13 round off = 1e-7 step 22*