Calculation of ESR spectra and related Fokker–Planck forms by the use of the Lanczos algorithm. II. Criteria for truncation of basis sets and recursive steps utilizing conjugate gradients[a)]

Kashyap V. Vasavada,[b)] David J. Schneider, and Jack H. Freed
Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853

(Received 14 July 1986; accepted 29 September 1986)

The complex symmetric Lanczos algorithm (LA) has proven to be a very efficient means of calculating magnetic resonance line shapes and spectral densities associated with Fokker–Planck forms. However, the relative importance of the various components of the basis set in an accurate representation of the spectrum and the proper number of recursive steps are not easily assessed in practice using the Lanczos algorithm. A systematic and objective procedure for the determination of optimal basis sets and number of recursive steps is developed using a generalization of the conjugate gradient method (CGM) appropriate for the type of complex symmetric matrices occuring in these problems. The relative importance of the individual basis vectors is determined by using the CGM to obtain the "solution vector" from the set of algebraic equations defining the spectrum. This is done at several values of the sweep variable (e.g., the frequency or the magnetic field). The maximum (over these values of sweep variable) for each component of the solution vector is taken to be a measure of the overall importance of the corresponding basis vector in the complete spectrum. Using this method signficant basis set truncation is conveniently possible. The number of recursive steps needed for an accurate representation of the spectrum is easily obtained by monitoring the residual in the approximate solution vector at the center of the spectrum and by recognizing the close relationship between the LA and the CGM. It is this relationship that enables construction of the Lanczos tridiagonal matrix with the CGM which can either be used to calculate the cw ESR spectrum directly or else the eigenvalues. The information obtained from the CGM can be used to "turbocharge" the LA by taking advantage of the nearly optimal basis set and number of recursive steps. Significant savings in computation time are possible, and relative savings are greatest for the most difficult problems. This is illustrated with a variety of examples of slow-motional cw ESR spectra and of the new two-dimensional electron-spin-echo technique. In keeping with the greater sensitivity of the latter technique to motional dynamics, it is consistently found to require significantly larger optimal basis sets and number of recursive steps for an accurate representation. One of the most challenging problems for both types of spectroscopy is the case of macroscopically oriented samples where the macroscopic director is tilted at an angle relative to the applied static magnetic field, since this removes much of the symmetry in the problem. This case is found to yield to very significant truncation of basis sets, and a new symmetry-based decoupling of certain basis vectors was found in this study for the particular example of a 90° tilt angle.

I. INTRODUCTION

The stochastic-Liouville theory[1] has provided a very useful framework for the analysis of ESR spectra. In this formalism, the intensity of absorption $I(\Delta\omega)$ is given by

$$I(\Delta\omega) = (1/\pi)\ \mathrm{Re}\langle v|[i(\Delta\omega\mathbf{1} - \mathbf{L}) + \mathbf{\Gamma}]^{-1}|v\rangle\,, \quad (1)$$

where $\Delta\omega$ is "the sweep variable", $\mathbf{L}$ is the Liouville operator associated with the spin Hamiltonian of the spin probe, and $\mathbf{\Gamma}$ is the diffusion operator for the stochastic variables that modulate the magnetic interactions. Also $|v\rangle$ is the so-called "starting vector" constructed from the spin transition moment averaged over the equilibrium ensemble. The vectors and operators are defined in the direct product space of the ESR transitions and of the functions of the stochastic variables. (In ESR we usually have $\Delta\omega = \omega - \omega_0$ where ω_0 is the Larmour frequency at the center of the spectrum and ω is the angular frequency of the applied radiation field.) We may rewrite Eq. (1) as

$$I(\Delta\omega) = (1/\pi)\ \mathrm{Re}\langle v|u(\Delta\omega)\rangle\,, \quad (2)$$

where $|u(\Delta\omega)\rangle$ is the solution of the equation

$$\mathbf{A}'|u\rangle = |v\rangle\,. \quad (3)$$

The matrix $\mathbf{A}'$ is defined as $\mathbf{A}' = i\Delta\omega\mathbf{1} + \mathbf{A}$ where $\mathbf{A} = \mathbf{\Gamma} - i\mathbf{L}$. Equation (1) can be solved either by inversion of $\mathbf{A}'(\Delta\omega)$ for a range of values of $\Delta\omega$, or alternatively by diagonalizing $\mathbf{A}$ only once.[1]

Recently a new form of ESR spectroscopy, based upon the electron-spin-echo technique, has been developed.[2] It is a

[a)] Supported by NSF Grants CHE83-19826, DMR86-04200, and by NIH Grant GM25862

[b)] NIH Senior Fellow (1985–1986). On leave from: Dept. of Physics, Indiana–Purdue Univ., Indianapolis, Ind. 46223.

 0021-9606/87/020647-15$02.10

two-dimensional technique, which we call 2D-ESE. This technique is found to be more sensitive than conventional ESR to motional dynamics.[2] Although it requires expressions that are more complex than Eq. (1), once the matrix **A** is diagonalized, the 2D-ESE spectrum can readily be obtained.[2] However, the increased sensitivity is reflected in the challenging computational problem to simulate 2D-ESE spectra.[2] In the present work we primarily address the accurate and efficient calculation of both types of ESR spectra.

The matrix **A** is in general very large and sparse. The conventional methods[1] for solving Eq. (3) by inversion or by diagonalizing **A** prove to be too cumbersome. One soon runs out of memory even on mainframe computers, and the solution requires prohibitive amounts of computer time. To remedy this situation the Lanczos algorithm has been developed for complex-symmetric matrices, since **A** is typically of this form.[3,4] It is an efficient method for tridiagonalizing **A** and is particularly suited to the solution of large-sparse matrices. It has previously been shown that it can lead to *at least* order of magnitude reductions in computation time, and it yields results to the solution of Eq. (1) to a high degree of accuracy.[3,5] It was also pointed out that this Lanczos algorithm is appropriate for the general class of Fokker–Planck equations which can be represented by a complex symmetric Fokker–Planck operator **A**. For such cases, Eq. (1) would be associated with the spectral density, which is the Fourier transform of the time correlation function of a dynamical variable $v(t)$.[3,4] In fact, more generally, it was possible to establish the close connection between the Lanczos algorithm based upon a scheme of projection operators in Hilbert space, and the Mori projection scheme in statistical mechanics.[4,6] Thus, while applications to ESR spectroscopy will be our primary concern in this work, they may be regarded as prototypical of a wide range of applications in chemical physics.

There are several advantages of employing the Lanczos algorithm (LA) for sparse matrices. These advantages include the facts that: (i) only the nonzero matrix elements need to be stored, (ii) they are not altered during the operations, and (iii) they are the only ones involved in the matrix operations. Another advantage is due to the fact that the LA tridiagonalization proceeds by recursive steps or projections. If we let N be the dimension of the matrix, and n_s the number of recursive steps needed to converge to an accurate spectrum, then in all cases we have studied we find $n_s \ll N$. This inequality becomes more dramatic the more complicated the problem. In this sense, the Lanczos projections rapidly seek out, from an initial finite subspace of dimension N [which is spanned by the starting basis set of orthonormal vectors: $|f_j\rangle, j = 1,...,N$], a smaller subspace spanned by the Lanczos vectors (i.e., the basis vectors for the tridiagonalized form of **A**, or $\mathbf{T}_n$) written as $|\Phi_k\rangle$, $k = 1,...,n$. When $n = n_s$ these Lanczos vectors are a sufficient basis for accurately representing the spectrum. We refer to this by noting that the LA constructs subspaces that progressively approximate the "optimal reduced space" for the problem. These subspaces, spanned by the Lanczos vectors are related to Krylov subspaces,[7,8] and are generated from the sequence $\mathbf{A}^{k-1}|v\rangle$ for $k = 1,...,n$. (That is, the n dimensional Krylov subspace is spanned by the n linearly independent vectors $\mathbf{A}^{k-1}|v\rangle$.) Thus, the choice of $|v\rangle$ as the "starting vector" biases the projections in favor of this "optimal reduced space." It is easy to show that this Krylov subspace can only contain eigenvectors of **A** with a nonzero component along $|v\rangle$. For computing cw spectra from Eq. (1), a continued fraction method[3,4] can be used directly on $\mathbf{T}_n$, since $|v\rangle$ is itself the first Lanczos vector. In general, the time required for the LA tridiagonalization goes approximately as $n_s N(2n_E + 21)$, where n_E is the average number of nonzero matrix elements in a row of **A**.[3]

There have appeared, since I,[3] reports of computational methods for calculating ESR spectra based on Padé approximants[9] and on the Mori method,[10] which may be expected to be formally equivalent to the application of the LA.[3,4,6] This matter has recently been studied in detail by Dammers,[11] who finds that while all these methods are indeed formally equivalent, the LA is the most stable and efficient from a computational viewpoint.

There are, however, disadvantages to the use of the Lanczos algorithm for numerical applications. Its main weakness is its loss of orthogonality among the Lanczos vectors $|\Phi_k\rangle$ which it generates from the Krylov set $\mathbf{A}^{k-1}|v\rangle$ by a kind of Schmidt orthogonalization. This is due to accumulation of numerical round-off errors. As a result the Lanczos steps can, in practice, be continued beyond the original matrix dimension (i.e., it is possible to have $n > N$). This leads to repeated eigenvalues as well as to spurious eigenvalues (due to introducing Lanczos vectors not contained in the rigorous Krylov subspace defined by **A** and $|v\rangle$). In general, the ESR spectra are determined by only a small subset of eigenvectors, in particular those associated with eigenvalues (a_i) with weakest damping (i.e, smallest Re a_i), and approximations to these eigenvalues (or "clusters of eigenvalues", cf. I) are rapidly generated such that $n_s \ll N$ for convergence to the spectrum, well before round-off error can accumulate to the point where it can significantly affect the computation. However, round-off error can become a problem if one works with an unnecessarily large basis set N and/or performs too many Lanczos projections, n in the interest of guaranteeing convergence to the correct spectrum. Such matters are more of a problem for 2D-ESE, since they require significantly larger values of N and n_s than cw-ESR to achieve convergence, as we show below.

Another limitation of the Lanczos algorithm as developed in I, is the lack of a convenient and objective criterion for determining n_s. One typically calculates the spectrum repeatedly for a sequence of Lanczos steps until convergence is confirmed. This is time consuming, and also it can ultimately lead to substantial accumulation of round-off error as n becomes large.

Finally we note the truncation problem: the ESR spectra can be calculated to a good approximation by finite matrices of large enough dimension N; one wishes to truncate the space so as to minimize N consistent with the accurate computation of the spectra. This is referred to in I as the minimum truncation scheme (MTS). Knowledge of the MTS can greatly speed up calculations. In I simple empirical rules were established for some of the simpler ESR spectra

(i.e., those of comparatively small dimension). These had to be obtained by trial-and-error calculation of spectra with different basis sets in order to specify which types of basis vectors were important. This scheme is very time consuming as well as incomplete. In actual practice one tends to work with sets of basis vectors significantly larger than the MTS, since the latter has not been convenient to determine. This problem is exacerbated for 2D-ESE, since the MTS is already very large.

Thus it would be desirable to have convenient and objective criteria to determine first the MTS and then the minimum Lanczos steps, n_s consistent with it. This would (1) greatly reduce computation time; (2) guarantee that the results have converged; and (3) minimize potential difficulties from computer round-off error.

It is in this context that we explore the method of conjugate gradients (CG) for the solution of Eq. (3). This method was first proposed by Hestenes and Steifel[12] for solving linear equations, and practical conjugate gradient procedures exist for symmetric positive definite matrices $\mathbf{A}'$.[7] In fact, for symmetric positive definite matrices it is possible to establish a formal equivalence between the CG method in solving Eq. (3) and the LA.[7,8] In this work we shall consider this equivalence for the *complex* symmetric matrices that arise in the ESR (and general Fokker–Planck) problems, and then use the equivalence in a practical algorithm which can (1) tridiagonalize $\mathbf{A}_N$ (the N-dimensional representation of $\mathbf{A}$), (2) estimate at each recursive step the magnitude of an appropriately chosen residual, so that the recursive steps can be terminated when $n = n_s$, and (3) establish the MTS for the problem. This is possible by utilizing the CG method because it can directly solve for $|u\rangle$, and it can also be used to tridiagonalize $\mathbf{A}_N$ through its connection to the LA while it simultaneously provides an estimate of the residual. Thus the CG method can be applied to turbocharge the LA by first determining the MTS and the n_s for typical values of the relevant physical parameters, and then in subsequent calculations, the turbocharged LA can be employed. This can greatly speed up the comparison of theoretical to experimental spectra in order to fit these physical parameters.

In Sec. II we review the LA and the CG method and consider their equivalence. We discuss their applicability for the relevant complex symmetric matrices, and we develop the algorithm, which provides the residual at each recursive step. We also outline in this section our procedure for determining the MTS. In Sec. III we present the results of our computer studies on calculating cw ESR and 2D-ESE spectra including new rules that emerge in estimating the MTS. A summary appears in Sec. IV.

II. LA AND CG FOR COMPLEX SYMMETRIC MATRICES

A. The Lanczos algorithm

We briefly summarize the Lanczos algorithm, which is known to be an application of the method of moments in Hilbert space. We first identify the starting vector $|v\rangle$ as the first Lanczos vector $|\Phi_1\rangle$. Then a kind of Schmidt orthogonalization on the Krylov sequence $\mathbf{A}^{k-1}|v\rangle$ for $k = 1,...,n$ allows one to iteratively generate the set of orthonormal Lanczos vectors $|\Phi_k\rangle$ according to

$$\beta_{k+1}|\Phi_{k+1}\rangle = (1-\mathbf{P}_k)\mathbf{A}|\Phi_k\rangle\,, \tag{4}$$

where β_{k+1} is the normalizing coefficient such that

$$\langle\Phi_{k+1}|\Phi_{k+1}\rangle = 1 \tag{5}$$

and $\mathbf{P}_k$ is the projection operator on the Krylov subspace spanned by the $|\Phi_j\rangle$ given by[13]

$$\mathbf{P}_k = \sum_{j=1}^{k} |\Phi_j\rangle\langle\Phi_j|, \quad k \leqslant n\,. \tag{6}$$

Equation (4) leads to a three term recursive relation for generating the $|\Phi_j\rangle$:

$$\beta_{k+1}|\Phi_{k+1}\rangle = (\mathbf{A}-\alpha_k)|\Phi_k\rangle - \beta_k|\Phi_{k-1}\rangle\,, \tag{7}$$

where

$$\alpha_k = \langle\Phi_k|\mathbf{A}|\Phi_k\rangle \tag{8}$$

and

$$\beta_k = \langle\Phi_k|\mathbf{A}|\Phi_{k-1}\rangle\,. \tag{9}$$

It may easily be shown that $\mathbf{A}$ has a tridiagonal representation, $\mathbf{T}_n$ in the basis of Lanczos vectors $|\Phi_j\rangle$ such that

$$\langle\Phi_k|\mathbf{A}|\Phi_j\rangle = 0 \quad \text{if} \quad k \neq j, j\mp 1 \tag{10}$$

while Eqs. (8) and (9) give the nonzero matrix elements. That is, given the vectors $|\Phi_k\rangle$ in terms of their components $x_{j,k}$ in the original basis set, $|f_j\rangle, j = 1,...,N$,

$$|\Phi_k\rangle = \sum_j x_{j,k}|f_j\rangle\,, \tag{11a}$$

$$x_{j,k} = \langle f_j|\Phi_k\rangle\,, \tag{11b}$$

then the column vectors $\mathbf{x}_k$ form the orthogonal matrix $\mathbf{Q}_n$ such that $\mathbf{Q}_n^{\mathrm{tr}}\mathbf{Q}_n = \mathbf{1}_n$ and

$$\mathbf{T}_n = \mathbf{Q}_n^{\mathrm{tr}}\mathbf{A}_N\mathbf{Q}_n\,. \tag{12}$$

We have described the conventional Lanczos algorithm for real symmetric (or Hermitian) matrices $\mathbf{A}$ such that Eq. (5) involves the usual norms in Hilbert space.

For our present applications to ESR (and Fokker–Planck equations) for which $\mathbf{A}$ is complex symmetric (or else can be transformed to complex symmetric form[3,4,6]) Moro and Freed[3,4] showed that one must introduce the Euclidean pseudonorm. That is, first consider the general non-Hermitian case. One must introduce a biorthonormal set of functions Φ_j and Φ^j such that

$$\langle\Phi^j|\Phi_j\rangle = \delta_{j,j'} \tag{13}$$

or alternatively (letting $\mathbf{x}^j$ and $\mathbf{x}_j$ be their column vector representations):

$$(\mathbf{x}^j)^{\dagger}\cdot\mathbf{x}_j = \delta_{j,j'}\,. \tag{14}$$

However, for the case of (nondefective)[8] complex symmetric matrices $\mathbf{A}$, it is possible to let

$$\mathbf{x}^j = \mathbf{x}_j^* \tag{15}$$

such that Eq. (14) becomes

$$\mathbf{x}_j^{\mathrm{tr}}\cdot\mathbf{x}_j = \delta_{j,j'} \tag{16}$$

and then the recursion method of Eqs. (4)–(10) remains applicable with Eq. (16) defining the Euclidean pseudonorm, whereby the bra vectors are defined without the usual complex conjugation in a Hilbert space.

Finally we note that the complex symmetric tridiagonal matrix $\mathbf{T}_n$ can easily be diagonalized by one of several methods.[8,14] While this is not necessary for cw ESR spectra (or for simple spectral densities from Fokker–Planck equations), it is needed for 2D-ESE spectra.

B. The conjugate gradient method

The conjugate gradient method of Hestenes and Steifel[12] is for solving equations of form $\mathbf{A}|u\rangle = |v\rangle$ with $\mathbf{A}$ a real symmetric positive definite (RSPD) matrix. The starting point of the CG method is to consider this equation in the form

$$|r_k\rangle = |v\rangle - \mathbf{A}|u_k\rangle\,, \tag{17}$$

where $|u_k\rangle$ is the kth approximant to $|u\rangle$, and $|r_k\rangle$ is the residual vector associated with $|u_k\rangle$. [This $|r_k\rangle$ is easily seen[7] to be the vector which gives the negative gradient of the functional $f(u_k) = \frac{1}{2}\langle u_k|\mathbf{A}|u_k\rangle - \langle u_k|v\rangle$ provided $\mathbf{A}$ is RSPD, so that a minimization of $f(u_k)$ is equivalent to solving $\mathbf{A}|u\rangle = |v\rangle$.] Equation (17) is solved by successive iterations which do not minimize along $|r_k\rangle$ for $k = 1,...,n$ (that would be the method of steepest descent), but instead minimize along a set of "conjugate directions" $|p_k\rangle$ $k = 1,...,n$ (which keeps the successive minimizations from spoiling those along previous conjugate directions) such that

$$|r_{k+1}\rangle = |r_k\rangle - a_k\,\mathbf{A}|p_k\rangle \tag{18}$$

and

$$|p_{k+1}\rangle = |r_{k+1}\rangle + b_k|p_k\rangle\,, \tag{19}$$

where the a_k and b_k are given by

$$a_k = \langle r_k|r_k\rangle/\langle p_k|\mathbf{A}|p_k\rangle\,, \tag{20}$$

$$b_k = \langle r_{k+1}|r_{k+1}\rangle/\langle r_k|r_k\rangle\,, \tag{21}$$

and where the $|r_k\rangle$ are easily shown to be mutually orthogonal (but they are not normalized), while the set of $|p_k\rangle$ are "A conjugate," i.e.,

$$\langle p_k|\mathbf{A}|p_j\rangle = 0 \quad \text{if } j \neq k\,. \tag{22}$$

The $|p_k\rangle$ is the closest vector to $|r_k\rangle$ that is A conjugate to the $|p_1\rangle \cdots |p_{k-1}\rangle$.[15] It is also true that $\langle p_j|r_k\rangle = 0$ for $j = 1,...,k-1$. Also, the $k+1$ approximant $|u_{k+1}\rangle$ is obtained from $|u_k\rangle$ as

$$|u_{k+1}\rangle = |u_k\rangle + a_k|p_k\rangle\,. \tag{23}$$

These equations permit one to iteratively obtain the higher order approximants. That is, let

$$|r_1\rangle = |v\rangle - \mathbf{A}|u_1\rangle \tag{24}$$

and

$$|p_1\rangle = |r_1\rangle\,, \tag{25}$$

where $|u_1\rangle$ is the initial guess of $|u\rangle$, then for $k = 1,...,n-1$ one calculates successively a_k, $|u_{k+1}\rangle$, $|r_{k+1}\rangle$, b_k, and $|p_{k+1}\rangle$ according to Eqs. (18), (19), (20), (21), and (23).[16] At each step the norm of $|r_k\rangle$:

$$\|r_k\|^2 = \langle r_k|r_k\rangle \tag{26}$$

is a measure of the extent of convergence to the final solution $|u\rangle$.

At this stage the LA and CG appear to be very different algorithms, the former tridiagonalizes $\mathbf{A}$ (or $\mathbf{A}'$), while the latter solves Eq. (3) (with $\Delta\omega = 0$) iteratively for $|u\rangle$. We shall discuss their equivalence below. For now we consider the applicability of this CG method to complex-symmetric matrices. The standard CG method requires that[7,8,17] $\mathbf{A}$ be RSPD. Nevertheless, one may first note that the CG method is a special case of the method of conjugate directions.[17] It is not hard to show, from this more general approach that for nondefective (i.e., diagonalizable), and nonsingular complex symmetric matrices, the above CG method applies provided only that we use the Euclidean pseudonorm [e.g., Eq. (16)] in our Hilbert space. This is just as we found for application of the LA to complex symmetric matrices, and the analysis is similar. However, for the CG method, we do have the additional requirement that $\mathbf{A}$ or $\mathbf{A}'$ be nonsingular. Let us consider replacing $\mathbf{A}$ by $\mathbf{A}'$ in the above equations recalling (cf. Sec. I) that the complex symmetric matrix $\mathbf{A}$ will have complex roots: the real parts give the damping and the imaginary parts the resonance frequencies. Since all the relevant eigenmodes of $\mathbf{A}$ will experience some damping, then, in general, both $\mathbf{A}$ and $\mathbf{A}'$ will be nonsingular and Eqs. (17)–(26) can apply for either matrix. Nevertheless, it is convenient to replace $i\Delta\omega$ by $i\Delta\omega + T_2'^{-1}$ where $T_2'^{-1}$ is a very small width parameter (cf. below).

There is, however, only one place where we relax the use of Euclidean pseudonorm: that is in the norm of $|r_k\rangle$, given by Eq. (26), which is used to estimate the extent of convergence. Given $|r_k\rangle$ that is determined from Eq. (18), then we have considered two specific forms of the norm, Eq. (26) which we write in terms of its components $y_{j,k}$ in the original basis set [cf. Eqs. (11)]:

$$\text{(i)}\ r_{k,ps}^2 = \left|\sum_j y_{kj}^2\right|\,, \tag{27}$$

$$\text{(ii)}\ r_{k,H}^2 = \sum_j |y_{kj}|^2\,, \tag{28}$$

whereas letting $|r_{k,\text{true}}\rangle \equiv |v\rangle - \mathbf{A}'|u_k\rangle$ [cf. Eq. (17)] at each iterative step we have a third norm:

$$\text{(iii)}\ r_{k,\text{true}}^2 = \sum_j |y_{\text{true},k,j}|^2\,. \tag{29}$$

For a complex symmetric matrix, only the second and third norms are equal in exact arithmetic and are guaranteed to be real. The first norm is, however, consistent with the Euclidean pseudonorm required in our CG algorithm for complex symmetric matrices, but the (unnormalized) $|r_k\rangle$ vectors would yield a complex value for $r_{k,ps}^2$ unless the absolute value is specified as in Eq. (27). In practice, we do find that the norms (ii) and (iii) are equal in finite precision arithmetic to just about the limit of double precision accuracy ($\|r\| \approx 10^{-11}$ on the PRIME 9955 computer we use), while (i) is always smaller. Once this limit is reached, any further attempt to improve the calculation by iteration is, of course, unsuccessful, and $r_{k,\text{true}}^2$ remains constant as k is increased. However, the other two forms of $\|r\|$ [(i) and (ii)] continue to decrease and are of no further value. But (i) and (ii) are readily available during each iteration, while (iii) requires extra calculation. We have used $r_{k,H}^2$ as our criterion for convergence, so we shall henceforth call it r^2. On the other

hand, it is advisable to occasionally use $r^2_{k,\text{true}}$ to check if round-off has become a problem.

Thus, we have found that the CG method can be used to solve for $|u\rangle$ in Eq. (3) which then permits a solution of the ESR spectrum $I(\Delta\omega)$ from Eq. (1). We have gained a specific criterion, $r^2_{k,H}$ to determine when convergence has been achieved. However, in this form, one must solve $|u(\Delta\omega)\rangle$ for each value of the sweep variable, a decided disadvantage compared to the LA from which a single tridiagonalization permits $I(\Delta\omega)$ to be calculated by the continued fraction procedure.[3] The deficiency of CG is more serious for predicting 2D-ESE spectra, since one *must* obtain eigenvalues of A.

To proceed further with the CG method we must recognize its equivalence to the LA, a matter which is not generally appreciated by the scientific community. Before discussing this equivalence, we first wish to point out that the orthogonal set of vectors $|r_k\rangle$ and the conjugate set $|p_k\rangle$ are contained in the same Krylov subspace of $\mathbf{A}^{k-1}|v\rangle$, $k = 1,...,n$ as are the Lanczos vectors $|\Phi_k\rangle$.[17]

C. Equivalence of CG and LA

The equivalence of CG and LA is discussed in several places for the case of RSPD matrices A.[7,8] In particular Golub and Van Loan[7] give an explicit expression for the construction of the Lanczos tridiagonal matrix $\mathbf{T}_n$ by the CG method for a RSPD A which is valid at each stage of the iteration. They show that

$$\mathbf{T}_k = \mathbf{\Delta}_k^{-1}\,\mathbf{B}_k^{\text{tr}}\mathscr{A}_k\,\mathbf{B}_k\mathbf{\Delta}_k^{-1}, \tag{30}$$

where $\mathscr{A}_k$ is the diagonal matrix with elements:

$$\mathscr{A}_{k;ii} = \langle p_i|\mathbf{A}|p_i\rangle, \quad i = 1,...,k \tag{31}$$

and $\mathbf{\Delta}$ is also diagonal with elements:

$$\Delta_{k,ii} = \|r_i\| = \left(\sum_j y_{i,j}{}^2\right)^{1/2} \equiv \rho_i, \quad i = 1,...,k \tag{32}$$

while $\mathbf{B}_k$ is an upper bidiagonal matrix with elements

$$B_{k;ii} = 1, \quad B_{k;i,i+1} = -b_i, \quad i = 1,...,k \tag{33}$$

with b_i given by Eq. (21). Furthermore, the unnormalized residual vectors $|r_i\rangle$ are collinear with the orthonormal Lanczos vectors $|\Phi_i\rangle$. More precisely

$$|\Phi_i\rangle = \mp(\rho_i)^{-1}|r_i\rangle, \quad i = 1,...,n. \tag{34}$$

Since the signs of the $|r_i\rangle$ are well-defined by Eq. (17), we see that the Lanczos vectors, which are normalized to within an arbitrary $\mp$ sign (for Euclidean norms) bear the sign ambiguity.[18] It follows from Eqs. (30)–(33) that the α_k, the diagonal elements of $\mathbf{T}_n$, and the β_k, its off-diagonal elements, are given by

$$\alpha_k = \langle p_k|\mathbf{A}|p_k\rangle/\rho_k^2 + (\rho_k^2/\rho_{k-1}^4)\langle p_{k-1}|\mathbf{A}|p_{k-1}\rangle \tag{35}$$

and

$$\beta_k = -(\rho_k/\rho_{k-1}^3)\langle p_{k-1}|\mathbf{A}|p_{k-1}\rangle. \tag{36}$$

Thus the α_k and β_k are readily obtained by CG at each iteration, so this approach may be used to tridiagonalize A. Also one must start with $|r_1\rangle = |v\rangle$ which implies $|u_1\rangle = 0$ from Eq. (24), in order to obtain Eq. (34) and thus to achieve correspondence between CG and the LA.

We turn now to the applicability of this approach for nondefective, nonsingular, complex symmetric matrices. Again we find that it is appropriate for such cases provided we use the Euclidean pseudonorm. This means that ρ_i as defined by Eq. (32), is a complex quantity [we choose the principal value of the square root in Eq. (32)].

We have verified this equivalence to the LA for our complex-symmetric matrices A with the use of the Euclidean pseudonorm by numerical calculation. We find that the $\mathbf{T}_n$ obtained by the LA [Eqs. (8) and (9)] and that from CG [Eqs. (35) and (36)] agree numerically in magnitude to at least six significant figures. However, there are sign differences between the β_k's of Eqs. (9) and (36), which are attributable to the arbitrary sign in the normalization of the $|\Phi_i\rangle$ [cf. Eq. (34)]. Such sign differences do not matter in calculating $I(\Delta\omega)$ or in the eigenvalues.

We ran into one problem in our use of CG. Because of the structure of our matrix representation of A and of $|v\rangle$, a division by zero can occur in the first CG step. A simple remedy is to use a small but finite value of "intrinsic linewidth" in all diagonal entries (i.e., we add a constant matrix to A which is diagonal). It can simply be subtracted from all the eigenvalues, once they are computed, if desired. (In fact, an "intrinsic linewidth" is normally added after tridiagonalization or full-diagonalization.)

Thus we conclude that the CG procedure can be applied to our complex-symmetric matrices A to obtain the Lanczos tridiagonal matrix from which spectra may be calculated. In the CG method we have an objective criterion, Eq. (28), of the extent of convergence at each iteration. Finally, we note that CG can be used to directly solve the linear equation problem of Eq. (3) when desired. In fact, this will serve as the basis of our approach to the determination of the MTS, as we describe next.

D. Minimum truncation scheme

As we discussed in Sec. I, it would be highly desirable to have an objective and convenient criterion for selecting the minimum basis set for representing A, which still guarantees convergence to the desired accuracy. We make use of the CG method to calculate $|u(\Delta\omega)\rangle$ for different values of the sweep variable, $\Delta\omega$. Since the spectrum is determined by the scalar product $\langle v|u(\Delta\omega)\rangle$ [cf. Eq. (2)], then we expect that a knowledge of the vector $|u(\Delta\omega)\rangle$ in terms of its components $z_j \equiv \langle f_j|u(\Delta\omega)\rangle$ in the original basis set, $|f_j\rangle$ $j = 1,...,N$ [cf. Eq. (11)], for a sample set of sweep positions, would provide an accurate assessment of the importance of each basis vector $|f_j\rangle$ in determining $|u(\Delta\omega)\rangle$.

Consider the jth component z_j, then from Eq. (3):

$$z_j(\Delta\omega) = \langle f_j|\mathbf{A}'^{-1}|v\rangle = \sum_m \langle f_j|\psi_m\rangle a'(\Delta\omega)_m^{-1}\langle\psi_m|v\rangle, \tag{37}$$

where, in the last equality of Eq. (37) we have introduced the eigenvectors of A' [or A] as the set $|\psi_m\rangle$, and $a'(\Delta\omega)_m$ [or $a(\Delta\omega)_m$] are their eigenvalues. This last expression in Eq. (37) is a product of three quantities. First, the scalar product $\langle\psi_m|v\rangle$, which is the projection of the mth eigenvector on the transition moment vector, is a measure of the

importance of this eigenvector in contributing to the spectrum. Next, $\langle f_j | \psi_m \rangle$ is a measure of the importance of the jth basis vector in contributing to $|\psi_m\rangle$. Finally $a'_m(\Delta\omega)^{-1}$ expresses *how* the mth eigenvector contributes to the spectrum, e.g., if it resonates far from the applied rf field, or, if it is very broad, then $a'_m(\Delta\omega)^{-1}$ is very small. All these factors are needed to estimate the importance of $|f_j\rangle$ to the spectrum. Since they are all included in z_j, then $z_j(\Delta\omega)$ can be used as a measure of the importance of $|f_j\rangle$ to the spectrum, provided only that we sample it at a sufficient number of positions in the spectrum (i.e., for sufficient number of values of sweep variable $\Delta\omega$). This may be done by solving Eq. (3) by CG for these positions.

Our approach therefore requires that a basis set larger than the MTS, but containing the latter as a subset, be utilized initially. Nevertheless, in most applications, wherein calculations are compared to experimental spectra, it is necessary to vary theoretical parameters and to repeat the computation many times, so the initial efforts at selecting the MTS can often be worthwhile. As problems become larger, it is usually possible to estimate a starting basis set that is not excessively large by extrapolation from an empirical set of rules derived from the MTS obtained from smaller, but closely related problems.[19] Then the final search for the MTS for the larger problem is less time consuming. Our examples below, illustrate this.

Given that Eq. (37) is a good criterion for determining the MTS, an alternative way to proceed would be to diagonalize **A** by the LA but to store the transformation matrices to obtain the $\langle f_j | \psi_m \rangle$ and the $\langle \psi_m | v \rangle$. Unfortunately it would destroy the great efficiency of the LA to keep track of the full transformation matrix.[7,8] Consequently Cullum and Willoughby[8] recommend an inverse iteration procedure to obtain the eigenvectors once a set of converged eigenvalues have been obtained by the LA. However, the spectra are extremely well approximated well before the actual eigenvalues have converged.[3] Thus, much more effort is required in implementing the LA in order to achieve accurate enough eigenvalues to permit estimates of $\langle f_j | \psi_m \rangle$ by inverse iteration, than is required to obtain our converged spectra. (This is perhaps less true in the case of 2D-ESE spectroscopy). On the other hand, the basic CG algorithm successfully delivers the needed information to determine the MTS as discussed below.

III. COMPUTER RESULTS

A. Slow-motional ESR spectra

This is the application to which Moro and Freed in I first applied the LA. We have applied the CG method discussed in Secs. II B and II C for the calculation of slow-motional ESR spectra (cf. Fig. 1) to obtain the relevant tridiagonal matrices $\mathbf{T}_n$ and residuals r_0^2 (i.e., r^2 calculated for $\Delta\omega = 0$). The spectra given by Eq. (1) could then be solved directly by the continued-fraction approach used in I. There is only a marginal increase in computation time ($\sim 5\%$) for the CG algorithm vs our LA.

We find, in general, that r_0^2 behaves like a damped oscillator as a function of n, i.e., it oscillates, but its local averge value decreases as n increases. In general, the cw spectra converge very rapidly; typically a value of $r_0^2 \approx 10^{-3}$–10^{-4} is more than adequate to obtain a good cw spectrum. We recommend, as a conservative criterion, that the iterations be terminated when $r_0^2 \approx 10^{-4}$, even though for most of the spectra a surprisingly low value of $r_0^2 \approx 10^{-2}$ is sufficient. (A consideration of the relationship between r_0^2 and the actual error in the spectrum is given below.) We therefore take as n_s the value of n when r_0^2 first reaches a value of 10^{-4}.

The slow-motional ESR spectrum depends upon a number of theoretical parameters viz: the hyperfine and g tensors, the rotational diffusion tensor **R**, and, for oriented fluids (e.g., liquid crystals or membranes) one must specify the mean orienting potential as well as the orientation (ψ) of principal axis of macroscopic alignment (the director) relative to the applied static magnetic field.[1,20] For convenience in the present studies we have let **R** be isotropic with a rotational diffusion coefficient R, and we have let the dimensionless mean potential to be of the simple form: $-U(\theta)/kT = \lambda P_2(\cos\theta)$ [where $P_2(\cos\theta)$ is the Legendre polynomial of order 2 and θ is the angle between the principal molecular axis and the director]. Also, when $\lambda \neq 0$ we have let ψ be either 0° or 90°.

As the motion slows down (i.e., R decreases) one requires larger dimensional representations of **A** (i.e., increased N), and also there is an increase in n_s. This is true no matter which algorithm is utilized. In all cases $n_s \ll N$ in agreement with previous results. However, we do find that n_s increases more slowly than N, so the advantage of the LA or CG becomes relatively greater, the larger N becomes. Some typical results reflecting these features are shown in Table I. In obtaining the results in Table I we have used values of N significantly larger than the MTS to ensure convergence, as is usually done in applications of the LA.[2,20–23] However, in one case we compare with a smaller value of N that more closely corresponds to the MTS. In this case we find that n_s is only decreased by a factor of 1.34 when N is decreased by a factor of 4.1. This is an example of the more general phenomenon of how the LA seeks out an approximation to the "optimal reduced space" already discussed in I.

In the absence of an objective criterion of convergence from the LA itself, Moro and Freed in I utilized the following definition of the error in the cw spectrum as

$$\Delta_n = \int_{-\infty}^{\infty} |I_R(\omega) - I_n(\omega)| d\omega , \tag{38}$$

where I_R is the "exact absorption spectrum" (actually some very good approximation of it), and $I_n(\omega)$ is the spectrum obtained after n iterative steps. Both are normalized to unity. [In Eq. (38) we take ω as our sweep variable for simplicity.] Then n_s was defined as the minimum number of steps satisfying the condition $\Delta_n \leqslant 10^{-4}$, which is an accuracy better than can be achieved experimentally. We find from several examples [using as $I_R(\omega)$ the spectrum obtained for $r_0^2 = 10^{-10}$ and the same basis set used for $I_n(\omega)$] that an $r_0^2 \approx 10^{-4}$ corresponds to a $\Delta_n \approx 10^{-7}$ to 10^{-8} with the smaller Δ_n corresponding to smaller N (e.g., 429) and the larger for $N \approx 8000$. [For $r_0^2 \approx 10^{-2}$, $\Delta_n \approx 2.4 \times 10^{-5}$.] Note, however, that while Δ_n is a measure of the actual error in the spectrum, it does not serve as a practical indicator of convergence

TABLE I. Typical values of iterative steps required for convergence to a given residual for different basis sets.

N[a]	r_0^{2}[b]	n_s[c]	Case[d]	Matrix elements[e]
1743	10^{-2}	60	A[f]	32 917
	10^{-4}	104		
	10^{-10}	172		
429	10^{-2}	49	A[f]	7 701
	10^{-4}	77		
	10^{-10}	128		
3543	10^{-2}	76	B	70 399
	10^{-4}	159		
	10^{-10}	315		
7503	10^{-2}	89	C	288 085
	10^{-4}	170		
	10^{-10}	326		
8196	10^{-2}	57	D	667 965
	10^{-4}	80		
	10^{-10}	143		

[a] N is the dimension of the matrix defined by $L^{e}_{\max}$, $L^{o}_{\max}$, $K_{\max}$, and $M_{\max}$, and symmetries given in Refs. 1 and 20.

[b] r_0^2 is the residual squared [cf. Eq. (28)], calculated at the center of the spectrum.

[c] n_s is the number of CG iterations.

[d] Case A is for "Tempone" magnetic parameters (cf. Table II) and $R = 10^6\ s^{-1}$; case B is also for tempone parameters but $R = 10^4\ s^{-1}$; case C is for "CSL" parameters (cf. Table II) and $R = 10^4\ s^{-1}$. Case D is for tempone parameters but with a strong potential ($\lambda = 10$), $R = 10^6\ s^{-1}$, and director tilt $\psi = 90°$. Also $(\gamma_e\, T_2')^{-1} = 1$ G.

[e] Number of nonzero matrix elements for the full matrix A counting the real and imaginary parts separately.

[f] $N = 429$ corresponds to approximate MTS, cf. Table III entry 5, while $N = 1743$ represents a typical larger basis set commonly utilized in 2D-ESE simulations by LA.

during the calculation. One must already have a very good approximation to the spectrum to calculate Δ_n! On the other hand, r_0^2 is readily obtained, and, we find it serves as a useful measure.

A comparison of entries C and D in Table I brings out an important point. The matrix in case D has larger dimension and about 2.3 times as many matrix elements as those of case C. Yet to reach a value of $r_0^2 \approx 10^{-10}$ we need less than half the number of steps in case D than for case C. This shows that specification of r_0^2 is a much better criterion for convergence than specification of the number of recursive steps. For a new matrix it is not possible to guess correctly the value of n_s, while one can readily obtain r_0^2. Availability of r^2 at each step of iteration is a tremendous advantage of CG over LA.

One cautionary note is in order. Our calculation of r^2 for the above has been at the center of the nitroxide spectrum (i.e., r_0^2). In general, the results for r^2 will be a function of $\Delta\omega$ [cf. Eqs. (17) and (28)]. However, our above comparison between r_0^2 at the center of the spectrum and Δ_n, which measures deviations across the whole spectrum, shows that this is a sufficient test for the class of spectra considered. This is, perhaps, no surprise considering the success of the LA itself in these spectral problems. In other cases it could be necessary to choose $\Delta\omega$ in A′ to correspond to the region of poorest convergence in implementing CG, so that the r^2 test reflects such a region. (Alternatively computations at several values of $\Delta\omega$ might be needed to perform the CG calculation and to obtain the needed r^2 results.)

B. Two-dimensional ESE spectra (2D-ESE)

We have also used the CG method to study 2D-ESE spectra. We have used the approximate expression of Millhauser and Freed[2] to predict the signal:

$$S(\omega,\omega') \propto \sum_j c_j^2 \frac{T_{2j}}{1+\omega^2 T_{2,j}^2} \exp[-(\omega'-\omega_j)^2/\Delta^2]\,, \tag{39}$$

where for the jth "dynamic-spin packet" (i.e., the jth normal mode solution $|\psi_j\rangle$ to A corresponding to eigenvalue a_j), we have $T_{2,j}{}^{-1} = \mathrm{Re}(a_j)$ as its Lorentzian width and $\omega_j = \mathrm{Im}(a_j)$ its resonant frequency. Also

$$c_j^2 = \langle \psi_j|v\rangle^2 \approx \mathrm{Re}\langle \psi_j|v\rangle^2\,, \tag{40}$$

where the approximate equality is valid only in the very slow (i.e., for $R < 10^5\ s^{-1}$) motional region for which Eq. (39) is approximately valid. The 2D-ESE spectrum is inhomogeneously broadened (with respect to the ω' sweep variable) by convolution with a Gaussian distribution of half-width Δ. For purposes of testing the computational method we have utilized Eq. (39) with the approximate form for c_j^2 in Eq. (40) even when the motion is taken as too fast for these expressions to accurately represent the experiment. This enabled us to extrapolate convergence schemes to the very slow motional cases for which Eq. (39) is applicable before we performed the *very* time-consuming calculations required in these cases.

It is clear from Eq. (39) that it is necessary to obtain estimates of the eigenvalues a_j which contribute appreciably (i.e., for which $|\langle \psi_j|v\rangle|^2$ is not negligible). This was done by diagonalizing the tridiagonal matrix $\mathbf{T}_n$ utilizing standard procedures.[3,14] The $|\psi_j\rangle$ are then, in principle, obtained in terms of their components $\langle\Phi_k|\psi_j\rangle$ in the Lanczos basis set. However, only the components along $|\Phi_1\rangle = |v\rangle$ are needed, and they form a vector of dimension n_s, which is easily obtained during the procedure.[1,3,14]

Because the 2D-ESE spectra require significantly more accurate estimates of the eigenvalues a_j and the weights c_j^2, we find that convergence with respect to LA or CG steps occurs only after achieving a residual that is much smaller than what is required for the corresponding cw spectrum. In particular we find that $r_0^2 \approx 10^{-8}$–10^{-10} is sufficient for 2D-ESE spectra [by visual comparison of normalized contours (cf. Fig. 2) in several cases], whereas the cw spectra have already converged for $r_0^2 \geqslant 10^{-4}$. This much more severe requirement for r_0^2 fortunately does not require very many more iterations as illustrated by the results summarized in Table I.

Again we must comment that the r^2 test has been performed only for the center of the spectrum. However, our visual comparisons noted above indicate that it should be sufficient for nitroxide-type spectra.

In order to study the convergence further, we introduce, by analogy to Eq. (38) the following definition of the error in the 2D-ESE spectrum:

$$\Delta_{2D,n} = \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' |S_R(\omega,\omega') - S_n(\omega,\omega')| , \tag{41}$$

where $S_R(\omega,\omega')$ is the normalized "exact 2D-ESE spectrum" or some very good approximation to it, and $S_n(\omega,\omega')$ is the normalized spectrum obtained after n iterative steps. Both S_R and S_n are calculated from Eq. (39). We find from several examples that an $r_0^2 \approx 10^{-10}$ corresponds to a $\Delta_{2D,n} \approx 0.007$ to 0.02 [where we utilized for $S_R(\omega,\omega')$ an $r_0^2 = 10^{-12}$ and the same basis set].

Finally it should be noted that we have checked the CG eigenvalue method against the exact EISPACK routine (for complex general matrices) for the case where $N = 177$. The 2D normalized contours match exactly. As expected the EISPACK routine takes much more time than the CG method. The CPU times on the PRIME 9955 computer were 1118 and 15 s, respectively.

Before we discuss the MTS below, we wish to show that our use of Eq. (37) to estimate the importance of each basis vector $|f_j\rangle$ is reasonable for the 2D-ESE spectrum. We first set $\omega = 0$ in Eq. (39). Then using Eq. (40) we can rewrite:

$$S(0,\omega') \propto \sum_j \langle \psi_j|v\rangle \sum_k \langle v|f_k\rangle\langle f_k|\psi_j\rangle T_{2,j} \times \exp[-(\Delta\omega_j')^2/\Delta^2] , \tag{42}$$

where $\Delta\omega_j = \omega' - \omega_j$. Thus the contribution of $|f_k\rangle$ is given by

$$\langle v|f_k\rangle \sum_j \langle f_k|\psi_j\rangle T_{2,j} \exp[-(\Delta\omega_j')^2/\Delta^2]\langle \psi_j|v\rangle \leqslant \langle v|f_k\rangle \sum_j \langle f_k|\psi_j\rangle [\mathrm{Re}\, a'(\Delta\omega')_j]^{-1}\langle \psi_j|v\rangle \tag{43}$$

since $[\mathrm{Re}\, a'(\Delta\omega')_j]^{-1} = T_{2,j}/[1 + (\Delta\omega_j' T_{2,j})^2]$ (i.e., a Lorentzian of width $T_{2,j}^{-1}$) which will be greater than the modified Gaussian expression: $T_{2,j} \exp[-(\Delta\omega_j')^2/\Delta^2]$, since in general, $\Delta \gg T_{2,j}^{-1}$ (i.e., relatively large inhomogeneous broadening). The inequality in Eq. (43) is to be "understood" in the sense of applying separately to the magnitudes of the real and the imaginary parts on each side. If we now assume $\langle \psi_j|v\rangle$ is essentially a real quantity according to Eq. (40), and recall (cf. above) that $\langle f_k|v\rangle$ is chosen real, it readily follows that $\langle \psi_j|f_k\rangle$ must be real. Thus the right-hand side of the above inequality can be rewritten as $\langle v|f_k\rangle \times \mathrm{Re}\, z_k(\Delta\omega')$. It thus follows from the discussion in Sec. II D that $z_k(\Delta\omega')$ [or more precisely $\mathrm{Re}\, z_k(\Delta\omega')$] may be used to estimate the importance of $|f_k\rangle$ to the 2D-ESE spectrum.

C. Minimum truncation schemes

1. Basis set

We have used as our basis set, $|f_j\rangle$ the standard set utilized for these problems in the high-field approximation.[1,3,20] They may be characterized by a set of five indices or quantum numbers as

$$|L,K,M,p,q\rangle = \left| \sqrt{\frac{2L+1}{8\pi^2}} D^L_{MK}(\Omega), p, q \right\rangle .$$

The $D^L_{MK}(\Omega)$ are the generalized spherical harmonics (or Wigner rotation-matrix elements) and Ω represents the Euler angles between the molecular frame and the lab frame, while M and K are $\geqslant 0$ and K is even for the cases studied here. They are the eigenfunctions of the rotational diffusion operator for an axially symmetric diffusion tensor $\mathbf{R}$ in an isotropic fluid. The quantum numbers p and q are defined as $p = m' - m''$ and $q = m' + m''$ where m' and m'' are the nuclear spin projection quantum numbers for the initial state and final states. Thus $p = 0$ ($p \neq 0$) for an allowed (forbidden) ESR transition. Our calculations were performed specifically for ^{14}N nitroxides with nuclear spin $I = 1$, so m', $m'' = -1, 0, +1$. Thus $p = 0, \mp 1, \mp 2$ and $q = -Q, -Q+2, \ldots, Q$ where $Q = 2 - |p|$. In isotropic fluids, and in ordered fluids where the tilt angle $\psi = 0°$ there is considerable symmetry of $\mathbf{A}$ such that $p = M$, so $M = 0$, 1, or 2. (The latter require also that the nuclear Zeeman term be neglected.) These symmetries help to reduce the basis set and contribute to the sparsity of $\mathbf{A}$. When $\psi \neq 0°$ the cylindrical symmetry in the lab frame is destroyed and the problem becomes more complicated.[20-22] We take the z axis in the lab frame to be along the director. This axis plus the direction of the static magnetic field (which is rotated relative to this z axis) define the x-z plane in the lab frame. For $\psi = 90°$ these axes become the lab z and x axes, respectively. In these cases the index M is no longer restricted, so p and M can take on their full range of values. Also the sparsity of $\mathbf{A}$ is reduced. As the motion slows down this requires immense basis sets, and $N \approx 10^3$–10^4 is possible. This problem is exacerbated for 2D-ESE given that larger basis sets are required for this more sensitive type of spectroscopy. Thus, it is especially important to minimize such basis sets in order to make the problems tractable.

2. Determining the significance of each basis vector

We have studied the problem of the MTS by utilizing the method discussed in Sec. II D, viz. the determination of the coefficients $z_j(\Delta\omega)$ for a particular set of values of the parameters (R,λ,etc.) by solving Eq. (3) with the CG method for several values of the sweep variable in the range of the spectrum (say $\Delta\omega/\gamma_e$ of -50 to $+50$ G about the center of the spectrum). We find that 10–20 are sufficient for slow-motional ESR spectra. In performing the sweep by CG it is useful to use as the initial vector [cf. Eq. (24)] at the mth sweep position $|u_1(\Delta\omega_m)\rangle$ the solution to the previous field position [i.e., $|u_n(\Delta\omega_{m-1})\rangle$] to accelerate the convergence. Also, it is sufficient for present purposes, which are just to estimate the importance of the $|f_j\rangle$, to use weaker convergence criteria for r^2 (e.g., $\approx 10^{-2}$ for cw spectra and $\approx 10^{-6}$ for 2D-ESE spectra). [Note that unlike the r_0^2 test used to terminate the recursions at n_s (cf. Secs. IIIA and IIIB), in the present case the r^2 test is applied to each spectral position calculated by CG.] One can get an idea of the computer time required from the following example. For a matrix of dimension $N = 1743$, the eigenvalue determination (with $r_0^2 = 10^{-10}$) took 215 s on a PRIME 9955 computer, while sweeping using 21 spectral positions (with $r^2 = 10^{-10}$) took 818 s after $\mathbf{A}$ had been preconditioned (cf. Sec. III D). In general, we find that solving for $|u\rangle$ at one spectral position is about five times faster on average than the diagonalization of $\mathbf{A}$. This is because signficantly fewer

iterations per spectral position are required when the initial approximate solution vector is the solution at the previous spectral position. Also, preconditioning speeds the convergence up somewhat (cf. Sec. III D).

The procedure we followed to estimate the significance of each basis vector $|f_j\rangle$ was to first determine the $z_j(\Delta\omega_i)$ $j=1,...,N$ at the ith spectral position. Then we computed the relative contribution to the spectrum $f_{j,i}$ as

$$f_{j,i} \equiv |z_j(\Delta\omega_i)|/|\langle v|u(\Delta\omega_i)\rangle| = |z_j(\Delta\omega_i)|\Big/\left|\sum_j x_{j,1} z_j(\Delta\omega_i)\right|, \quad i=1,...,m \quad (44)$$

[cf. Eqs. (2) and (11)]. This is repeated for all m spectral positions. As the spectral positions are scanned, only the maximum value for each $|f_j\rangle$ is stored, so at the end, one has a vector of N entries: $f_{j,\max}$, which give the global maxima for each $|f_j\rangle$. Then we determined which basis vectors have values of $f_{j,\max}$ (that we give in percentages) at or above the level of say 3%, 1%, 0.3%, etc. These were then used to simulate spectra. We were thus able to determine from a number of examples that retaining basis vectors for which $f_{j,\max} > 3\%$ would guarantee convergence for cw spectra. However, for the more sensitive 2D-ESE spectra, we found from our examples that it is necessary to retain basis vectors for which $f_{j,\max} > 0.06\%$ to 0.03%. We have examined these results in the light of the error criteria of Eqs. (38) and (41). We find that for entries 5 of Tables II and III [with an original basis set of $N=1743$ and $r_0^2 = 10^{-10}$ to determine $I_R(\omega)$ and $r_0^2 = 10^{-12}$ for $S_R(\omega,\omega')$ an $f_{j,\max} \approx 3\%$ yields a $\Delta_n \approx 3\times 10^{-4}$, thus corresponding closely to the criterion used in I for convergence of the spectrum. Also an $f_{j,\max} \approx 0.03\%$ yields a $\Delta_n \approx 4\times 10^{-8}$ and a $\Delta_{2D,n} \approx 0.01$.

We therefore can suggest as a conservative estimate of the MTS those basis vectors for which $f_{j,\max} > 3\%$ for cw ESR and $f_{j,\max} > 0.03\%$ for 2D-ESE. Of course, in preliminary simulations, one can use less stringent conditions to obtain rough approximations to the spectra.

In general, one expects that the convergence required depends on the degree of spectral resolution as represented by the extra or "intrinsic width" $T_2'^{-1}$. Its role can be inferred from Eq. (37) for the $z_j(\Delta\omega)$. An increase in this intrinsic width will decrease the magnitudes of the inverse eigenvalues $a'(\Delta\omega)_m^{-1}$ without affecting the other terms on the right-hand side. Thus, in general, fewer basis vectors would be needed if the same criteria for the MTS as given above are used. Also, it might be that, due to reduced spectral resolution resulting from a larger intrinsic width, the above criteria [obtained with $(\gamma_e T_2')^{-1} = 1$ G] could be more stringent than necessary.

3. General observations

We summarize in Tables II and III our respective results for cw-ESR and 2D-ESE in determining the MTS based upon these two criteria. In actual fact, we first used

TABLE II. Truncation parameters for cw spectra.

No.	Spin[a] probe	R[b] (s^{-1})	λ[c]	$L^e_{\max}$[d]	$L^o_{\max}$[d]	$K_{\max}$[d]	$M_{\max}$[d]	N[e]	N'[g]	$N_{\min}$[f]
1	Tempone	10^7	0	6	3	2	2	42	34	33
2	Tempone	10^7	10	10	none	2	2	63	26	26
3	Tempone (90° tilt)	10^7	1	6	3	2	6	288	134	74
4	Tempone (90° tilt)	10^7	10	10	9	4	4	822	129	69
5	Tempone	10^6	0	14	7	6	2	171	108	100
6	Tempone	10^6	5	12	3	2	2	78	54	42
7	Tempone	10^6	10	10	none	0	2	33	30	29
8	Tempone (90° tilt)	10^6	10	12	11	6	6	1779	533	245
9	Tempone	10^5	0	30	13	10	2	543	285	256
10	Tempone	10^4	0	54	15	10	2	990	549	447
11	CSL	10^6	0	14	7	14	2	231	174	162
12	CSL	10^5	0	30	13	30	2	762	522	474

[a] Values of g and A tensors are: Tempone: $g_{xx}=2.0088, g_{yy}=2.0061, g_{zz}=2.0027, A_{xx}=5.8, A_{yy}=5.8, A_{zz}=30.8$. CSL $g_{xx}=2.0021, g_{yy}=2.0089, g_{zz}=2.0058, A_{xx}=33.44, A_{yy}=5.27, A_{zz}=5.27$. Static magnetic field = 3300 G. $(\gamma_e T_2')^{-1} = 1$ G.

[b] Rotational diffusion coefficient.

[c] Potential strength parameter.

[d] $L^e_{\max}$, $L^o_{\max}$, $K_{\max}$, and $M_{\max}$ are the maximum values of L_{even}, L_{odd}, K, and M required at the level of $f_{j,\max} = 3\%$.

[e] N is the dimension of the matrix if *all* the states up to the maximum values of L^e, L^o, K, and M are included.

[f] $N_{\min}$ is the size of the minimum truncation scheme for basis vectors at the level of 3% or larger contribution.

[g] N' is the dimension of the matrix after the look-up table is applied to the N-dimensional basis set given in footnote e. This lookup table specifies an $M_{\min}$ and an $M_{\max}$ for each important pair of L and K. For $\psi = 90°$ the new selection rules discussed in the text were also utilized.

TABLE III. Truncation parameters for 2D-ESE spectra.[a]

No.	Spin probe	R (s^{-1})	λ	L^e_{max}	L^o_{max}	K_{max}	M_{max}	N	N'	N_{min}
1	Tempone	10^7	0	10	7	6	2	123	94	92
2	Tempone	10^7	10	16	7	2	2	108	81	76
3	Tempone (90° tilt)	10^7	1	10	7	6	10	1440	752	586
4	Tempone (90° tilt)	10^7	10	16	15	6	6	2601	931	607
5	Tempone	10^6	0	22	17	10	2	429	317	307
6	Tempone	10^6	5	20	15	8	2	333	223	209
7	Tempone	10^6	10	16	11	4	2	168	131	120
8	Tempone (90° tilt)	10^6	10	20	19	10	12	8196	3804	2835
9	Tempone	10^5	0	44	37	18	2	1485	1010	971
10	Tempone	10^4	0	88	71	28	2	4614	2706	2506
11	CSL	10^6	0	22	19	22	2	600	503	485
12	CSL	10^5	0	46	37	46	2	2310	1877	1815

[a] Parameters have the same meaning as in Table II. Unlike Table II, however, all states contributing at least 0.03% are included for the sensitivity required to simulate 2D-ESE spectra.

basis sets significantly larger than the MTS as has been the normal procedure.[20–23] Then we determined a simple set of truncation rules corresponding to the maximum values of indices: L^e, L^o, K, and M that are consistent with our results for the MTS yielding a basis set of dimension N, which is given in the Tables. [Note that we distinguish between even and odd L values by superscripts e and o.] The actual N_{min} corresponding to the MTS is always smaller than this.

Some general observations are that the values of N and N_{min} are consistently several times greater for 2D-ESE vs cw ESR as one might anticipate from the greater sensitivity of the former (i.e., $f_{j,max}$ of 0.03% vs 3%). In the cases for isotropic fluids or for ordered fluids with $\psi = 0°$ (i.e., no tilt), $N_{min} < N$, but N_{min} is at most about a factor of 2 smaller with the largest differences for the slowest motions. This implies that the simple truncation rules of defining an L^e_{max}, L^o_{max}, K_{max}, and M_{max} in Tables II and III yield most of the pruning of basis vectors for the relatively faster motions. Yet substantial further pruning can be realized in the cases of very slow motions. On the other hand, cases with $\psi = 90°$, which yield much larger matrices, are characterized by values of N_{min} that are much smaller than the values of N obtained by the simple truncation rules.

Our calculations, summarized in Tables II and III are for two classes of nitroxide problems. In the first class we assume that the hyperfine tensor **A** is axially symmetric, and we can choose the principal axis of diffusion to be parallel to the principal (cylindrical) axis of **A** so that $A_{zz} > A_{xx} \simeq A_{yy}$. This would be possible, for example, if the rotational diffusion tensor **R** were isotropic. Since this is frequently the case for the small TEMPONE (2,2,6,6-tetramethyl-4-piperidone-1-oxyl) probe, we shall nominally refer to the first class as TEMPONE parameters.

In the second class, the principal axis of an axially symmetric tensor **R** is taken to be along the principal x axis of the **A** tensor. Such would be the case for the long rigid probe CSL[21] (cholestane), which normally exhibits anisotropic diffusion, so we nominally refer to this class as CSL parameters. For both classes we consider the limiting case of isotropic R, since this usually implies the largest MTS. We see from Tables II and III that the choice of CSL parameters leads to larger MTS, as would be expected for the reduced symmetry. (Note that the g tensor used in these tables is not axially symmetric as in the case for nitroxides.)

4. Truncations in the absence of orienting potential

The best understood calculations are those for the cw ESR line shapes for the TEMPONE class. The maximum L, K, and M needed for an adequate representation of the cw ESR spectra of this class may be found in Tables II and III for a sequence of diffusion rates. The important trends to note are that L^e_{max} increases by a factor of 2 for every order of magnitude decrease in diffusional rate, and that L^o_{max} and K_{max} follow a similar pattern except they seem to remain almost constant for diffusion rates slower than $10^5\ s^{-1}$. Also, symmetry restrictions require $M_{max} \leqslant 2$.

The difference between N and N_{min} is due primarily to the fact that the $M = 1$ and 2 (therefore the singly and doubly forbidden spin transitions) are not important for large values of K, and the maximum value of K depends on the particular value of L. Thus, it would be insufficient to specify a single K_{max}, as in the Tables, to achieve a basis set of dimension close to N_{min}. The maximum value of K needed usually occurs for $L < L_{max}$, and the full tensorial basis is needed up to this point. The maximum K for $L > K_{max}$ then decreases as L increases. This phenomenon, together with the fact that the larger M values are not important for large K, suggest that the discrepancy between N and N_{min} could be significantly decreased if a maximum (and minimum value) for the M index for evey important value of K and L could be incorporated into the definition of the basis set. We therefore constructed a lookup table which lists each set of L and K

values that contributes one or more basis vectors to the MTS, and for each such set it lists the maximum and minimum value of M. The value of N', the dimension of the basis set determined with the aid of the lookup table is much closer to $N_{\min}$ than is N (cf. Tables II and III) indicating that the development of new truncation rules would be valuable even for these relatively simple problems.

The basis sets required for 2D-ESE spectra of the TEMPONE class are much larger but qualitatively quite similar in character to the basis sets needed for cw ESR calculations. The major difference in the pattern of important basis states is that $L^{o}_{\max}$ is closer to $L^{e}_{\max}$ for 2D-ESE calculations, and the fact that the $L^{e}_{\max}$ dependence is more along the lines expected from the results in I, which suggested an inverse cube root dependence of $L_{\max}$ on R. The pattern of diminished importance of the states with large M and K indices and states with large L and K seen for cw ESR is also observed here. (It should be noted that only the two slowest diffusion rates shown in the tables are slow enough for the simplified 2D-ESE theory to be strictly valid.) The ratio of N' to $N_{\min}$ is roughly the same for cw ESR and 2D-ESE.

The CSL class of spectra require a substantially larger basis set for the same diffusion rate as mentioned previously. A comparison of the basis sets for the two classes of probes show that $L^{e}_{\max}$ and $L^{o}_{\max}$ are nearly identical for both cw ESR and 2D-ESE spectra. In contrast, in all cases $K_{\max} = L^{e}_{\max}$ for CSL calculations while for TEMPONE, $K_{\max} \leqslant (1/2) L^{e}_{\max}$. The major qualitative differences between the basis sets needed for TEMPONE and CSL classes for the same diffusion rate are that for the latter, states with large L and *small* values of K are less important, whereas the $M \geqslant 1$ states are important for large K but not for small K. The 2D-ESE results for the CSL class of parameters show patterns similar to the cw ESR results. Note also the effectiveness of a lookup table defining an $M_{\min}$ and $M_{\max}$ for each L and K as demonstrated by the close agreement of N' and $N_{\min}$ for the CSL case.

For a case illustrated in Tables II and III we have plotted cw spectra and 2D contours. In the figures we exhibit two representative examples to show the type of spectra we are considering. In Figs. 1 and 2 the cw derivative spectrum (entry 10, Table II) and the 2D contours (entry 10, Table III) are shown, respectively.

5. Truncations in the presence of orienting potential but no director tilt

The introduction of a potential into the diffusion operator qualitatively changes the truncation problem because of the introduction of large off-diagonal real matrix elements. For very slow motions, a strong orienting potential can actually decrease $L^{e}_{\max}$, $L^{o}_{\max}$, and $K_{\max}$. Because of this phenomenon, the agreement between N and $N_{\min}$ is much better than in other slow-motional cases. The increase of $L^{e}_{\max}$ for $R = 10^{7}\,\mathrm{s}^{-1}$ appears consistent with earlier observations on cases of faster (but still slow) tumbling.[1]

6. Truncations in the presence of orienting potential and 90° director tilt

Calculations involving both an orienting potential and a nonzero director tilt were done only for the TEMPONE class. The trends of larger $L_{\max}$ and $K_{\max}$ for smaller R de-

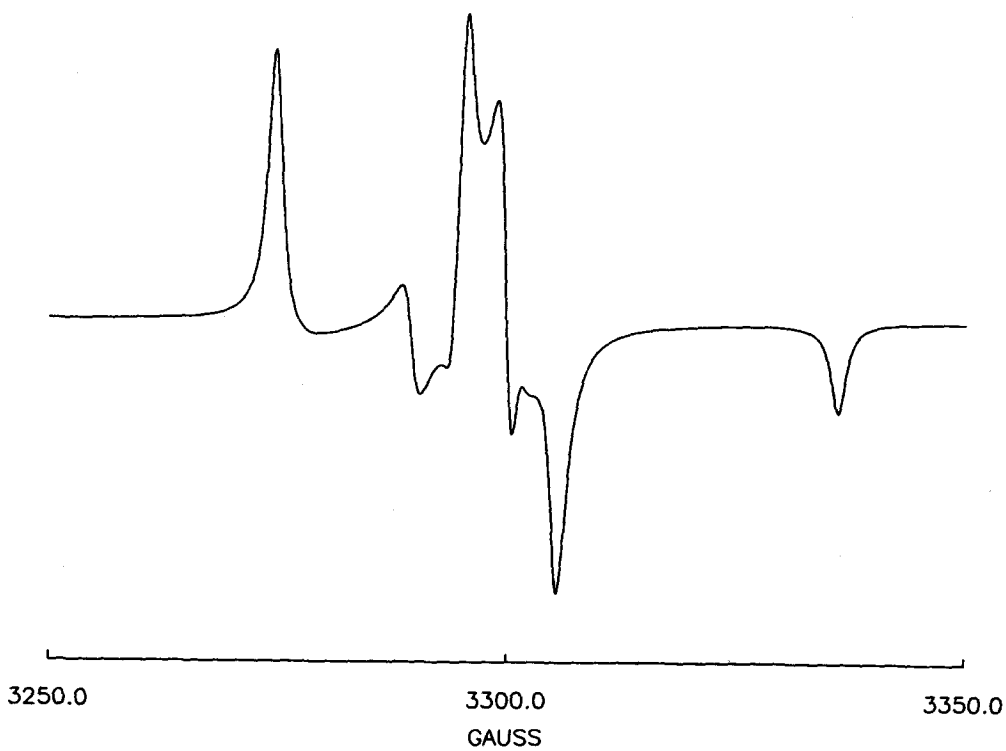

FIG. 1. cw ESR derivative spectrum. Parameters correspond to entry 10 in Table II. An "intrinsic linewidth" of 1 G has been added.

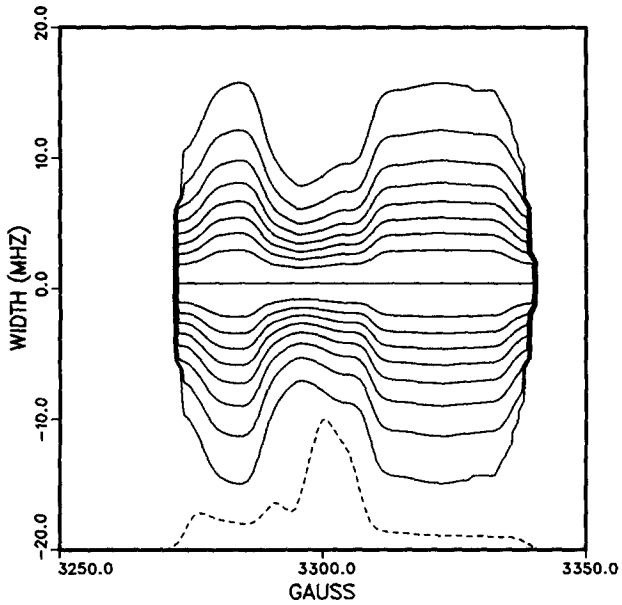

FIG. 2. Normalized 2D contours and the $\omega = 0$ slice obtained from 2D-ESE spectra (cf. Ref. 2). Parameters correspond to entry 10 in Table III. An "intrinsic linewidth" of 1 G has been added.

scribed previously are clearly evident here. In addition, the nonzero director tilt breaks the $M = p$ restriction, thus $M_{\max} > 2$ is allowed and is required to adequately represent the spectrum.

There are, however, some rather surprising facts that emerge from a detailed study of the important basis vectors. For instance, there are large classes of basis vectors which are not coupled to the starting vector. Since the starting vector for the case of a cylindrically symmetric potential contains nonzero elements only for states with even L and $K = M = p = 0$, only the states which are coupled to these by the stochastic Liouville operator are important. The only interaction term included in these calculations which can couple states with different M and p indices is the A (hyperfine) superoperator. However, because of the special symmetry of the $\psi = 90°$ problem, the states with

(1) even L,$M = 0$, $p = 1$, and $q = \pm 1$,

(2) odd L,$M = 0$, $p = 2$, and $q = 0$,

(3) all L, even M, $p = 0$, and $q = 0, \pm 2$

are found to be decoupled. This decoupling does not depend on the approximation of axial symmetry for A. It appears that this decoupling is a result of the properties of the Wigner rotation matrices, $d^2_{m,m'}(\psi)$, for $\psi = 90°$ which occur in the definition of the spin Liouville superoperator matrix elements. (Investigations on other symmetry-related simplifications of the 90° tilt problem are in progress.)

The large difference between N and $N_{\min}$ underscores the importance of further truncation. The introduction of the simple lookup table of the type described above improves the agreement, but clearly does not completely resolve the discrepancy. The remaining difference between N' and $N_{\min}$ is due to the fact that not all symmetry-allowed values of p and q are important for large values of L. The elimination of these states will require a more complete lookup table and/or more extensive truncation rules.

7. Comments

The estimation of the MTS is a very difficult problem to describe in general terms because of the large number of input parameters needed to define the problem. Only a fraction of the full range of these parameters has been considered here. It appears that there are three ways to proceed in defining a lookup table appropriate for a given problem; (1) to use a large database containing all the quantum numbers associated with each important basis vector for a wide range of input parameters; (2) to develop semiempirical rules, as illustrated above and in I, to handle the complicated trends observed in this work which involves more complicated cases than considered in I, and (3) a hybrid of the two previous methods.

The database approach is straightforward in many respects. One would simply generate lists of important quantum numbers for each given set of input parameters (and truncation criterion) and add these lists to the database. The items in the database could then be cross-referenced according to type of magnetic tensor parameters, diffusion constants, tilt angles, etc. To do a calculation, one would search through the database for a closely related problem that was done previously. The major problem with this approach is defining "closely related" and managing the very large amounts of data needed to make a useful database of this type. On the other hand, this approach would be very effective in cutting down computation time if a very closely related problem was previously catalogued. In fact, the spectra themselves could be included in such a database. This might be valuable in choosing an initial set of input parameters for the fitting of a new experimental result.

A second approach, based on the development of empirical rules to determine an appropriate $M_{\min}$ and $M_{\max}$ for each important L and K is also possible. In many ways this approach is more appealing, because it attempts to organize the very large amount of information along physically plausible lines. This approach suffers from the necessity of producing a complicated set of rules that could deal with the entire body of knowledge at once by finding useful patterns.

The third method, a hybrid, is based on the use of a database in the initial stages from which a reasonable set of empirical rules could be found. This would allow one to split up the problem into many smaller subproblems which could be studied individually. For instance, the importance of the individual elements of a block of states with a given L value could be studied as a function of input parameters. The patterns found for each L block could then be compared to give an overall scheme. The database could be assigned a secondary role once truncation rules are established and verified.[19]

D. Direct calculations of spectra and spectral densities by conjugate gradients

We have already employed CG to calculate the cw ESR spectrum for several values of the sweep variable in order to determine the MTS. It is, of course, possible to employ this direct approach to compute the entire cw spectrum from Eq.

(3). In our previous discussion (in Sec. III C) we estimated that the calculation at one field position is about five times faster on average than the diagonalization of A. This would mean that for a complete spectrum, for which typically 200 values of the sweep variable would suffice, the direct method would take 40 times longer.

One might, however, hope that as the CG calculation is performed for smaller increments of the sweep variable, then even fewer iterations per sweep position would be required when the initial guess is the solution to the previous sweep position. We have performed some numerical experiments for a matrix of $N = 1743$ to test this. For example we find that if we increase the number of sweep positions from 21 to 201 (i.e., by a factor of 9.57) the computer time required to complete the calculation increases by a factor of 5.47 (for $r^2 = 10^{-4}$ termination) and by a factor of 3.61 (for $r^2 = 10^{-2}$ termination). This indicates that one does improve the efficiency of the calculation in this manner especially when a larger r^2 is used. It should be mentioned, however, that if a larger r^2 (e.g., 10^{-2}) is used for field sweeping, some smoothing procedure and/or extrapolation methods (cf. below) will be necessary. Based upon our numerical experiments we expect that further efficiency could be achieved by letting the increment in sweep variable to be variable so that calculations are performed with smaller (larger) increment in regions where the spectrum is rapidly (slowly) changing.

Another way in which we have speeded up the direct calculation (in the above) is by preconditioning the $\mathbf{A}'$ matrix. Preconditioning is a device to improve the convergence of an iterative solution to a matrix problem [e.g., Eq. (3)]. Given a symmetric positive definite matrix $\mathbf{A}' = \mathbf{M} - \mathbf{N}$, one finds that the CG method can be accelerated utilizing $\mathbf{M}$ as a preconditioner provided $\mathbf{M}$ is symmetric and positive definite.[7] In our case of complex symmetric matrices, if we let $\mathbf{M} = \mathbf{\Gamma} + T_2'^{-1}\mathbf{1}$ and $\mathbf{N} = i(\mathbf{L} - \Delta\omega\mathbf{1})$, then $\mathbf{M}$ can be made symmetric and is positive definite. The preconditioned problem one solves is

$$\tilde{\mathbf{A}}'(\mathbf{M}^{1/2}|u\rangle) = \mathbf{M}^{-1/2}|v\rangle\,, \tag{45a}$$

where the preconditioned matrix $\tilde{\mathbf{A}}'$ is

$$\tilde{\mathbf{A}}' = \mathbf{M}^{-1/2}\mathbf{A}'\mathbf{M}^{-1/2}\,. \tag{45b}$$

For isotropic rotational diffusion, where the eigenvalues of the diffusion operator are proportional to $L(L+1)$, the stochastic-Liouville matrix becomes dominated by the real parts of the diagonal elements for large L. The effect of the preconditioning is to set these real parts equal to unity for all diagonal elements and to scale the other matrix elements accordingly. For more general cases (including a potential), in order to keep the calculation of $\mathbf{M}^{-1/2}$ simple we have just let $\mathbf{M}$ be diagonal and containing just the diagonal elements of $\mathbf{\Gamma} + T_2'^{-1}\mathbf{1}$.

In general, we do find that the preconditioned CG algorithm[7] does speed up the convergence of the calculation. For example, a case with $N = 1743$ required 104 CG steps to reach $r^2 = 10^{-4}$ before preconditioning, but after preconditioning it required only 45 steps. Unfortunately, because preconditioning is not a similarity transformation, it cannot be used to diagonalize $\mathbf{A}$ for purposes of calculating spectra. We employed the faster preconditioned algorithm for the results reported in Tables II and III.

Based upon these results we believe that tridiagonalization by the turbo-LA is the more efficient method for cw ESR, and we recall that the direct method does not even apply to 2D-ESE. However, there are some cases where diagonalization methods are not so suitable. These include (1) ESR on transition metal ions which require a wide range of sweep of the static magnetic field,[24] and (2) ESR in the presence of strong saturating radiation fields.[25,26] In the former case the contribution of the g tensor in the stochastic-Liouville equation (SLE) (which is dependent on the static magnetic field) is itself varying considerably, so one would have to diagonalize the spectrum for many different values of the static magnetic field. In the latter case, it is no longer possible to factor out the sweep variable from $\mathbf{A}'$ as a constant matrix as was done in Sec. I. For such cases the direct method by CG should be useful.

Further enhancement of the speed and efficiency of the direct CG method can be made by acceleration of the convergence by extrapolation methods. We may expect that only the first several recursive steps of CG actually provide new information; later steps merely remove undesirable "transients." An effective means of improving the convergence rate of a sequence is the Shanks transformation.[27] For the present case, let us define the spectrum obtained after n CG steps at sweep position $\Delta\omega_i$ from Eq. (2) by $I_n^{(0)}(\omega_i)$. Then the Shanks transformation in its simplest form would provide a new approximation $I_n^{(1)}(\omega_i)$ given by

$$I_n^{(m+1)}(\omega_i) = \frac{I_{n+1}^{(m)}(\omega_i) I_{n-1}^{(m)}(\omega_i) - I_n^{(m)}(\omega_i)^2}{I_{n+1}^{(m)}(\omega_i) + I_{n-1}^{(m)}(\omega_i) - 2I_n^{(m)}(\omega_i)} \tag{46}$$

for $m = 0$. An $(m+1)$th order approximant can be obtained by iterating this expression $m+1$ times. Equation (46) can be generalized to a k th-order Shanks transform by employing $2k+1$ members of the sequence $I_{n-k}^{(m)}(\omega_i)$, $I_{n-k+1}^{(m)}(\omega_i)$,..., $I_{n+k}^{(m)}(\omega_i)$.[27,28] We have made a preliminary study of the generalized and iterated Shanks transforms, and we find they are quite effective in accurately estimating $I(\omega_i)$ utilizing only $n \lesssim 20$, in some cases. Thus, while the full power of such methods has not been fully explored, they do show substantial promise.

IV. SUMMARY

We have found that the conjugate gradient method (CGM) may readily be extended to the complex symmetric matrices found in ESR spectral problems. These examples are also prototypical of Fokker–Planck forms. While the CGM can be used to calculate spectra by solving for the spectrum at each value of the sweep variable, its lesser known equivalence to the LA leads to an algorithm which provides the Lanczos tridiagonal form from which cw spectra can be computed by continued fractions, while 2D-ESE spectra can be computed after diagonalizing the tridiagonal matrix. This CGM also provides the residual r_0^2 at each recursive step, which serves as a "built-in" objective criterion to determine when convergence to the desired accuracy has

been achieved. Such a convenient criterion is lacking in the conventional LA. We find that an $r_0^2 \sim 10^{-4}$ is a conservative estimate of convergence for cw ESR spectra, but an $r_0^2 \sim 10^{-10}$ is appropriate for 2D-ESE spectra, thereby also emphasizing the greater sensitivity of the latter type of spectroscopy.

The same CGM, but utilized in the format for calculating spectra, may be applied at several key spectral positions to directly determine the importance of each original basis vector in contributing to the spectrum at each of these positions. A simple criterion of relative importance may then be utilized to determine the minimum truncation scheme (MTS). We estimate from our studies that a basis vector, which contributes at least 3% at any cw-spectral position be included in the MTS, whereas for 2D-ESE it is advisable to retain basis vectors which contribute at least 0.03%, again emphasizing the greater sensitivity of the latter experiment.

In past applications of the LA it was usually necessary to choose significantly larger basis sets than the MTS or alternatively to repeat a calculation with basis sets of different size in order to guarantee convergence because of the absence of an objective procedure to determine the MTS. We have found from a variety of examples of cw ESR and 2D-ESE spectra that simple rules for truncation may be developed utilizing our method. Such rules facilitate the selection of basis sets that are reasonable approximations to the MTS. This is especially true for ESR problems without an orienting potential, or with a potential but where the macroscopic director associated with this orienting potential is aligned parallel to the static magnetic field **B** to preserve the cylindrical symmetry in the lab frame. When this symmetry is destroyed by tilting the director at an angle ψ with respect to **B**, then, in general, many more basis vectors are required, and the dimension of the problem is vastly increased. This is one of the most challenging problems in the simulation of ESR spectra. An important observation that we have made in this work, is that such cases can be severely truncated beyond those provided by the simple rules useful for the $\psi = 0°$ case. Based upon our objective determinations of the MTS for cases where $\psi = 90°$, we determined a new set of truncation rules that could not be forseen given the complicated structure of the SLE when $\psi \neq 0°$. We even found a new selection rule applicable to the case of $\psi = 90°$, which had not previously been appreciated. As a result, the "tilt" case of $\psi = 90°$ is now rendered a much more tractable problem, and future work along these lines should lead to the MTS for the range of $0° < \psi < 90°$. The fact that the matrix structure of the SLE becomes so complicated probably means that the implementation of the MTS would require a combination of simple truncation rules along with a computer-generated "lookup" table with a listing of the relevant basis vectors.

Once the MTS and the required number of recursive steps have been determined for a given range of the relevant physical parameters it is possible to greatly speed up further calculations by the LA. This is expected to be particularly useful in fitting experimental results to these parameters by (nonlinear) least squares methods which necessarily involve many computations of the spectrum in the course of a single fitting.

Finally we note that a combination of selection of the MTS, plus the use of r_0^2 to determine the needed number of Lanczos recursions, is expected to provide, in complicated problems, more reliable calculations that will not be masked by round-off error, since (1) the r_0^2 test can be used to determine when round-off error has become a problem, and (2) only the minimum number of floating point arithmetic operations will be required once the MTS is used.

In a more fundamental sense, we recall from previous work the close connection between the Lanczos projection operator and the Mori projection operator in statistical mechanics.[4,6,13] Our analysis of the CGM and its correspondence with the LA also appears to lead to a further clarification of the relation between these methods and that of Mori.[15]

These above features can be expected to make the turbo-LA approach an extremely efficient procedure for calculating ESR (and NMR) spectra as well as other Fokker–Planck forms.

ACKNOWLEDGMENTS

We wish to thank Professor C. Van Loan and Mr. Larry Fried for useful discussions.

[1] J. H. Freed, in *Spin Labeling: Theory and Applications*, edited by L. J. Berliner (Academic, New York 1976), Vol. 1, Chap. 3 and references therein.

[2] G. L. Millhauser and J. H. Freed, J. Chem. Phys. **81**, 37 (1984); G. L. Millhauser, Ph. D. Thesis, Cornell University, 1986.

[3] G. Moro and J. H. Freed, J. Chem. Phys. **74**, 3757 (1981). Hereafter referred to as I.

[4] G. Moro and J. H. Freed, in *Large-Scale Eigenvalue Problems*, edited by J. Cullum and R. Willoughby, Mathematics Studies Series, Vol. 127 (North–Holland, Amsterdam, 1986).

[5] G. Moro and J. H. Freed, J. Phys. Chem. **84**, 2837 (1980).

[6] G. Moro and J. H. Freed, J. Chem. Phys. **75**, 3157 (1981).

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations* (The Johns Hopkins University, Baltimore, Maryland, 1983).

[8] J. K. Cullum and R. A. Willoughby, *Lanczos-Algorithms for Large Symmetric Eigenvalue Computations* (Birkhauser, Boston, 1985), Vol. I.

[9] A. J. Dammers, Y. K. Levine, and J. A. Tjon, Chem. Phys. Lett. **88**, 198 (1982).

[10] M. Giordano, P. Grigolini, D. Leporini, and P. Marin, Phys. Rev. A **28**, 2474 (1983).

[11] A. J. Dammers, Ph. D. Thesis, University of Utrecht, 1985.

[12] M. R. Hestenes and E. Steifel, J. Res. Natl. Bur. Stand. **49**, 409 (1952).

[13] Note that the Mori [Prog. Theor. Phys. **34**, 399 (1965)] projection operator $\mathscr{P}_j$ is closely related to P_k of the method of moments. This relation is expressed as $P_k = \Sigma_{j=1}^{k} \mathscr{P}_j$.

[14] R. G. Gordon and T. Messenger, in *Electron-Spin Relaxation in Liquids*, edited by L. T. Muus and P. W. Atkins (Plenum, New York, 1972).

[15] In Mori's treatment (cf. Ref. 13) he introduces the orthogonal set of functions f_k and complementary functions $\dot{f}_k \equiv iL_k f_k$ [where $L_k = (1 - \Sigma_{j=1}^{k} \mathscr{P}_j)\mathbf{L}$ with L the Liouville operator]. These f_k's are collinear with our $|\Phi_k\rangle$ and hence with the unnormalized $|r_k\rangle$. The $\dot{f}_k = i\omega_k f_k + f_{k+1}$, and thus they appear to bear a close relationship to the $\mathbf{A}|p_k\rangle$ given by Eq. (18), which we rewrite in the form $d_k \mathbf{A}|p_k\rangle = (\|r_k\|/\|r_{k+1}\|)|\Phi_k\rangle + |\Phi_{k+1}\rangle$ with d_k a coefficient.

[16] This is essentially algorithm 10.2-13 in Ref. 7 which is essentially that of Ref. 12.

[17] V. V. Voyevodin, *Linear Algebra* (Mir, Moscow, 1983).

[18] In Ref. 7 another algorithm [9.3-1] for a symmetric-positive definite matrix is developed to solve the *linear equation* problem by a blend of LA and CG. It defines the relevant vectors to avoid such a sign ambiguity. It also makes clear that if one starts with the LA algorithm, additional computations, equivalent to what is already done in CG, are required to obtain the residuals.

[19] The use of methods of artificial intelligence [cf. J. Chang and R. E. Wyatt,

Chem. Phys. Lett. **121**, 307 (1985)] could be useful in this context.
[20]E. Meirovitch, D. Igner, E. Igner, G. Moro, and J. H. Freed, J. Chem. Phys. **77**, 3915 (1982).
[21]E. Meirovitch, A. Nayeem, and J. H. Freed, J. Phys. Chem. **88**, 3454 (1984); E. Meirovitch and J. H. Freed, *ibid.* **88**, 4995 (1984).
[22]L. Kar, E. Ney-Igner, and J. H. Freed, Biophys. J. **48**, 569 (1985).
[23]L. Kar, G. L. Millhauser, and J. H. Freed, J. Phys. Chem. **88**, 3951 (1984).
[24]R. F. Campbell and J. H. Freed, J. Phys. Chem. **84**, 2668 (1980).
[25]S. A. Goldman, G. V. Bruno, and J. H. Freed, J. Chem. Phys. **59**, 3071 (1973).
[26]J. S. Hyde and L. Dalton, in *Spin Labeling: Theory and Applications*, edited by L. J. Berliner (Academic, New York, 1979), Vol. 2.
[27]C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw–Hill, New York, 1978), Chap. 8.
[28]Such generalized Shanks transforms are closely related to Padé approximants. If $k<n$, then the k th order Shanks transform is identical to the Padé approximant P_k^n of the series[27] $I_1^{(m)}(\omega_i)+\Sigma_{j=1}^{\infty}[I_{j+1}^{(m)}(\omega_i)-I_j^{(m)}(\omega_i)]z^j$, evaluated at $z=1$. [The LA and CG methods can themselves be shown to be equivalent to Padé approximants (cf. I and Ref. 11).]