*
From: LASERS, MOLECULES AND METHODS

Edited by:

Joseph O. Hirschfelder
Robert E. Wyatt
Rob. D. Coalson

Vol. 73 in Advances in Chemical Physics,
(1989).

# CHAPTER X *

# SPIN RELAXATION AND MOTIONAL DYNAMICS

D. J. SCHNEIDER and J. H. FREED

*Department of Chemistry, Cornell University, Ithaca, New York 14853*

## CONTENTS

I. Introduction
II. ESR Lineshapes and the Stochastic Liouville Equation
  A. Derivation of the Spectral Lineshape Function in the Linear Response Regime
  B. The Spin Hamiltonian
  C. Model Diffusion Operators
III. Overview of Discretization Methods for the Solution of the SLE
IV. Overview of Traditional Matrix Methods for the Solution of the SLE
  A. Solving Linear Systems of Equations Using Gaussian Elimination
  B. Complete Diagonalization Methods
    1. Jacobi-Type Methods
    2. Tridiagonalization and the *QR* Algorithm
V. Lanczos and Conjugate Gradients Methods of Solving the SLE
  A. The Lanczos Algorithm
  B. The Conjugate Gradients Algorithm
  C. The Equivalence of the LA and CG Methods
  D. Minimum Truncation Scheme
  E. Convergence of Lanczos–Conjugate Gradients Projections
  F. Calculating Two-Dimensional ESE Spectra
  G. Direct Calculation of Spectra and Spectral Densities by Conjugate Gradients
  H. Slow-Motional ESR Spectra: Examples
VI. The Fokker–Planck Approach to Modeling Molecular Dynamics
  A. General Fokker–Planck Operators
  B. Time Reversal, Detailed Balance, and Symmetrized Fokker–Planck Operators
    1. Reduction of General Fokker–Planck Operators to Complex Symmetric Matrix Form
    2. Calculation of Classical Time Correlation Functions and Spectral Densities with Complex Symmetric Lanczos Algorithm
    3. Example: The Planar Rotator

We wish to acknowledge support for this work by NSF grants DMR-86-04200 and CHE 87-03014 and NIH grant GM-25862 and by the Cornell University Materials Science Center.

C. Extensions of the Fokker–Planck Approach
D. Reduction of the Stochastic Liouville Operator to Complex Symmetric Matrix Form
VII. Nonlinear Phenomena: Saturation, Double Resonance, and Spin Echoes
A. The Stochastic Liouville Equation in the Presence of Radiation
B. General Methods of Solution
C. Steady-State Saturation and Double Resonance
D. Spin Echoes and Two-Dimensional ESE
E. Stimulated Echoes, Magnetization Transfer, and Two-Dimensional Fourier Transform Spectroscopy
VIII. Conclusions and Future Directions

## APPENDICES

A. Lanczos Algorithm: A Simple Derivation
B. Continued Fractions, Padé Approximants, and the Lanczos Algorithm
1. Continued Fraction Approximants Derived from the Lanczos Algorithm
2. Spectral Functions as Moment Problems
3. Padé Approximants to Spectral Functions
4. Summary
C. Bilinearly Metric Spaces and Relaxation Phenomena
D. Matrix Elements in ESR Problems
References

## I. INTRODUCTION

The study of molecular dynamics in liquids is an active and exciting area in theoretical, computational, and experimental chemical physics. The majority of experimental techniques for studying molecular dynamics in isotropic liquids and liquid crystalline phases involve measuring the response of the system to an externally applied time-dependent perturbation, usually electromagnetic radiation. Examples of such techniques are:

magnetic resonance (electron and nuclear spin resonance),
far-infrared and infrared absorption,
dielectric relaxation,
light scattering (Raman, Rayleigh–Brillouin, etc.),
fluorescence depolarization, and
inelastic neutron scattering.

The raw data from these experiments usually reflect the underlying molecular dynamics in a rather indirect fashion. In many instances it is necessary to explicitly model the response of the system to the perturbing field to extract quantitative information on the molecular motions that modulate the experimentally observed signals. The "inversion" of the spectroscopic data to expose the dynamical information can be a difficult task, especially if the perturbing field is so strong that the system cannot respond linearly to it.



In the weak perturbation limit, linear response theory can be used to simplify the interpretation of time domain experimental data by the connection between the observed signal and an equilibrium-averaged correlation function of the relevant dynamical variables. Likewise, frequency domain measurements in the weak perturbation limit reflect the spectral density of fluctuations in these dynamical variables at equilibrium. These two types of measurements are related by the fact that the spectral density observed in a frequency domain experiment is just the Fourier–Laplace transform of the correlation function obtained in a related time domain experiment. This relationship, a consequence of the fluctuation–dissipation theorem, is exemplified by the well-known equivalence between frequency-domain [continuous-wave (CW)] magnetic resonance spectra and free induction decay signals observed in the time domain. The study of molecular dynamics in liquids by modeling CW electron spin resonance (ESR) spectra in the linear response regime is a central topic of this review.

In more intense fields, where linear response theory is not applicable, the full equations of motion for the system including its coupling to the perturbing field must be solved. This is, in general, a more challenging computational problem, but it can lead to entirely new and informative types of spectroscopic tools to study molecular dynamics. In this review we also include a survey of CW nonlinear methods, such as saturation and double resonance, as well as newer ESR time domain multiple-pulse spin echo methods. These latter methods, when performed in an idealized manner, permit one to separately treat the effects of the intense radiation field and of the molecular dynamics; yet they allow for a great variety of intriguing possibilities in exploring the latter.

The various techniques for studying molecular dynamics in liquids can be roughly divided into two classes depending on the nature of the dynamical variables required to model the spectrum: (i) those that are sensitive to single particle properties and (ii) those that are sensitive to collective motions. For example, the Rayleigh–Brillouin light-scattering spectrum from a pure monatomic fluid such as argon depends on fluctuations in the electric polarizability density over distances on the order of the wavelength of visible light. These fluctuations surely involve the collective motion of many molecules. On the other hand, the ESR spectra of dilute solutions of spin probes in liquids are almost always interpretable in terms of the independent sum of the magnetizations from the individual molecules. The distinction between these two classes can become blurred. For instance, in the study of phase transitions in liquid crystals by ESR, the collective fluctuations in the

density or order parameter of the bulk fluid can couple to the spin degrees of freedom and dramatically affect the ESR spectrum.

In addition, the experiments can be classified according to the time scale over which they are most sensitive to fluctuations in the relevant variables. The time scale of a magnetic resonance experiment is determined, in part, by the magnitude of the fluctuations in the spin Hamiltonian produced by the interactions of the spin-bearing molecules with the surrounding solvent and by the spectral resolution. To clarify this statement, we must distinguish two limiting regimes of spin relaxation and two distinct types of relaxation. In the more familiar limit, the motional narrowing limit, one speaks of $T_1$ and $T_2$ types of spin relaxation. Now, the effectiveness of the $T_1$, or spin-lattice, relaxation depends in part on the relative magnitudes of the correlation rate (the inverse of the correlation time) of the dynamical fluctuations and the irradiating frequency. But it also depends on the ratio of the magnitude of the stochastic perturbations of the spin Hamiltonian to the correlation rate. However, the $T_2$, or transverse, relaxation depends only on the latter ratio. Nevertheless, even when this ratio is small, high-resolution spectroscopy enables the accurate measurement of small contributions to $T_2$ (e.g., to the CW linewidth). Thus, it is possible to detect very fast processes (correlation rates on the order of $10^{12}\,s^{-1}$), albeit in an indirect manner.

The other limit is the slow-motional limit. The slow-motional regime is where the stochastic perturbations to the spin Hamiltonian are comparable to, or greater than, the correlation rate. In this limit, there is no longer a time scale separation between the characteristic times of molecular motion and spin relaxation, which greatly simplifies the analysis of fast-motional spectra. Instead, the spin degrees of freedom and molecular dynamics become intimately coupled. In this limit, spin relaxation experiments probe the relaxation of these coupled modes and thus provide more direct information on the molecular dynamics. In conventional CWESR experiments with nitroxide spin probes, this limit is usually reached for correlation rates on the order of $10^9\,s^{-1}$, whereas for nuclear magnetic resonance (NMR), it can be on the order of $10^3\,s^{-1}$. The characteristic time scale for slow-motional ESR experiments implies that they will be sensitive to motions of spin probes in viscous liquids and/or slowly diffusing spin-labeled macromolecules in solution. On the other hand, the slower time scale of the typical NMR experiment is better suited to the slower dynamics in solid phases. Another feature of the magnetic resonance experiment is that, at least in principle, one can "tune in" the slow-motional regime by properly choosing the value of the magnetic field and the corresponding resonance frequency. This feature allows one to vary the relative magnitudes of some of the relevant terms in the fluctuating part of the spin Hamiltonian to increase the sensitivity of the experiment to a particular time scale of molecular motion. At present, this idea



is being used to extend the time scale of the slow-motional ESR experiment to the range $10^{10}$–$10^{11}\,s^{-1}$, depending on the nature of the spin probe, by working at high fields (90 kG) and high frequencies (250 GHz).

Whereas the analysis of slow-motional spectra holds the potential for greater information on molecular dynamics in liquids, it poses much more complicated computational challenges. In answer to these challenges, there have been a number of major advances in computational methods in recent years. These new computational methods for calculating slow-motional magnetic resonance spectra are the focal point of this review.

With these considerations in mind, the interpretation of dynamical effects on ESR spectra of dilute solutions of spin probes in viscous liquids can proceed in the following manner. Since the experimental observable depends only on the sum of the magnetization due to the individual spins and not on any collective phenomena or interaction between spins on separate particles, it makes sense to base the analysis on an approximate equation of motion for the one-particle spin density matrix. The time evolution of the macroscopic magnetization can then be calculated as the equilibrium average of the contributions due to the individual spins. Moreover, the nature of the interaction of the unpaired electron spin with its environment in a dilute solution is such that the spin Hamiltonian, which drives the time evolution of the density matrix, should depend only on the orientation and/or angular velocity of the spin-bearing molecule with respect to a laboratory-fixed reference frame and on the relative positions of pairs of spin-bearing molecules. This suggests that the positional, orientational, and angular velocity degrees of freedom of the spin probe can be modeled as a stochastic process in which the intricate details of the collisions with other molecules are unimportant.

In this manner, the molecular dynamics is incorporated into the calculation by assuming a specific form of the stochastic modulation of the orientation and position of the spin probe molecules, which can, in general, couple to other degrees of freedom of the fluid. It turns out to be particularly convenient and useful to model the molecular dynamics as a Markov process that modulates the various terms in the spin Hamiltonian. This is not, as we will see, a particularly restrictive approach. One is free to incorporate as many relevant dynamical variables as justified by the experiment into a multidimensional Markov process. However, the greater the number of degrees of freedom that are included, the more challenging is the computational problem. In this review we address the twofold problem presented by this stochastic modeling approach: (i) how to choose appropriate Markovian forms based on the known or presumed molecular physics of the system and (ii) how to solve the resulting equations for the relevant spectral densities or magnetic resonance lineshape. In the latter case, the spin dynamics and the molecular dynamics

can be described in a single equation of motion for the spin density matrix, which has been properly generalized to include its dependence on the orientation, angular velocity, and position of the molecule. This equation is usually referred to as the stochastic Liouville equation (SLE). The generalized spin density matrix combines the properties of the usual spin density matrix with those of a classical probability distribution for the dynamical variables incorporated into the Markovian model for the molecular dynamics. The calculation of magnetic resonance spectra from the SLE is the canonical problem dealt with in this review.

After we describe the basic ESR lineshape problem in terms of the SLE in Section II, we review various methods in Section III for reducing the set of coupled partial differential equations represented by the SLE to a tractable set of linear algebraic equations that can be solved on a computer. A survey of methods of solving the SLE based on the classic algorithms of numerical linear algebra are the subject of Section IV. We then focus in Section V on the newer methods based on the Lanczos [1]* and conjugate gradients algorithms [2] that have proven to be extremely powerful for these applications. We describe in some detail their strengths and how they may be employed.

There is a close theoretical connection between these methods and those used by other workers in a variety of fields. The theoretical interrelationships which provide a framework for understanding and justifying these methods are considered in Appendix B.

The modeling of the subclass of stochastic processes that can be described by Fokker–Planck equations is treated in Section VI. It is shown how the inherent symmetry of Fokker–Planck equations, which obey the requirements of detailed balance, allow one to extend the Lanczos and conjugate gradients methods in a particularly simple fashion to the calculation of spectral densities. In fact, what emerges is a very general approach to irreversible processes that obey the preceding restrictions. This general approach can be characterized in the following manner. Whereas the reversible dynamics (both classical and quantum mechanical) is most properly and conveniently analyzed using the familiar unitary or Hilbert space formalism, the inclusion of irreversible terms naturally leads to a formalism involving complex symmetric matrices and complex orthogonal spaces. Since the properties of complex symmetric matrices and nonunitary bilinearly metric spaces such as complex orthogonal spaces are unfamiliar to most readers, these ideas and their more familiar unitary space analogues are summarized in Appendix C because of their importance to the central themes of this review.

A brief survey of the influence of molecular dynamics on nonlinear phenomena and pulsed methods in magnetic resonance is given in Section VII.

*In this chapter, references are enclosed in brackets—Ed.



The complex symmetric Lanczos algorithm introduced by Moro and Freed [3] and the closely related complex symmetric conjugate gradients algorithm of Vasavada, Schneider, and Freed [4] arise in a natural fashion in the complex orthogonal space formalism for irreversible processes. A basic message of the present review is the applicability of these methods to computational and theoretical studies of irreversible processes in general and to the study of molecular dynamics in liquids by magnetic resonance in particular.

## II. ESR LINESHAPES AND THE STOCHASTIC LIOUVILLE EQUATION

The relationship between the ESR spectral lineshape function, $I(\Delta\omega)$ and the dynamics of motion of a paramagnetic molecule can be expressed in the form [5–12]

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

where $\Delta\omega$ is the sweep variable, $\mathbf{L}$ is the Liouville operator associated with the spin Hamiltonian of the probe molecule, and $\mathbf{\Gamma}$ is the Markovian operator for the stochastic variables that modulate the magnetic interactions. In most cases, $\mathbf{\Gamma}$ is taken to be a Fokker–Planck operator. 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 functions of the stochastic variables. (For typical ESR calculations $\Delta\omega = \omega - \omega_0$, where $\omega_0$ is the Larmor frequency at the center of the spectrum and $\omega$ is the angular frequency of the applied radiation field.) Equation 1 is derived from the more general stochastic Liouville equation, which is appropriate for studying the spectrum (cf. Section II.B).

By means of computer calculation of ESR spectra, one may extract information about the dynamics of motion. In particular, spectra in the so-called slow-motional region are sensitive to the form chosen for the diffusion operator $\mathbf{\Gamma}$, making it possible to distinguish between different models for the motion by comparison between experimental and theoretical spectra [12–25]. (The slow-motional regime may be defined by the inequality $|\mathbf{L}|/|\mathbf{\Gamma}| \geq 1$, where $|\mathbf{L}|$ and $|\mathbf{\Gamma}|$ are measures of the magnitude of matrix elements of $\mathbf{L}$ and $\mathbf{\Gamma}$, respectively.) This application of ESR spectroscopy depends on the efficiency of the algorithm for calculating spectra. As one utilizes more sophisticated models, calculations with matrices that increase geometrically in dimension are required. This need would seem warranted by the development

of new ESR techniques that are particularly sensitive to molecular motions [26–35].

Also, there are many applications of the ESR technique, requiring calculation of spectra, to systems of physical or biological relevance with the primary purpose of deriving the relaxation time(s) for the reorientational motion [12, 21–24, 36, 37]. In these applications, simple forms for the diffusion operator are utilized, so the size of the matrix is relatively small. But the use of a very compact algorithm that does not need large memory allows one to use a mini or personal computer directly connected to the ESR spectrometer. Also, one may obtain NMR lineshapes in the slow-motional region, particularly in solids, and these may also be analyzed in terms of Eq. 1 [36–40]. In fact, there continues to be a growing number of applications requiring detailed spectral calculations based on Eq. 1.

Our studies of ESR spectra and the modeling of motional dynamics have made clear that the same algorithms would be applicable to the general class of Fokker–Planck equations, since they may also be represented by operators of the form of those in Eq. 1 due to the existence of both inertial or drift terms and damping terms [3, 4, 41–44]. The calculation of the time correlation functions (or, more precisely, their Fourier–Laplace transforms, which are usually referred to as spectral densities) is also found to proceed from expressions like Eq. 1. Thus, the analysis and discussion in the next several sections will also apply to such cases.

Note that Eq. 1 can be rewritten as

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

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

$$\mathbf{A}'(\Delta\omega)|u(\Delta\omega)\rangle = (i\,\Delta\omega\,\mathbf{I} + \mathbf{A})|u(\Delta\omega)\rangle = |v\rangle. \tag{3}$$

The operator, $\mathbf{A}$ is defined as $\mathbf{A} = \mathbf{\Gamma} - i\mathbf{L}$. The spectrum given by Eq. 3 can be calculated by either solving Eq. 3 for a range of values of $\Delta\omega$ or, alternatively, by diagonalizing $\mathbf{A}$ only once [5–8, 12, 45].

The matrix of the operator $\mathbf{A}$ is in general very large and sparse. Thus, conventional methods [5–8, 12, 45] for solving Eq. 3 by inversion or by diagonalizing $\mathbf{A}$ prove to be too cumbersome (cf. Section IV). 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 (LA) has been developed for complex-symmetric matrices, since $\mathbf{A}$ is typically of this form or else it can be transformed to this form. It is an efficient method for tridiagonalizing $\mathbf{A}$ and is particularly well suited to the solution of sets of linear algebraic equations such as Eq. 3, which are characterized by large sparse matrices. We find 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, 4, 43, 44]. In a more theoretical vein, it was possible to establish the close connection between the LA based on a scheme of projection operators in Hilbert space and the Mori projection scheme in statistical mechanics [43, 44]. Though the emphasis here will be on applications to ESR spectroscopy, the Lanczos methods described in Section V may be regarded as appropriate for a wide range of applications in chemical physics.



There have appeared other reports of computational methods for calculating ESR spectra based on Padé approximants [46] and on the Mori method [47, 48], which may be expected to be formally equivalent to the application of the LA [3, 43, 44] (see also Appendices A and B). This matter has recently been studied in detail by Dammers [49], who finds that whereas all these methods are indeed formally equivalent, the LA is the most stable and efficient from a computational viewpoint. We discuss these matters in more detail in Appendix B.

### A. Derivation of the Spectral Lineshape Function in the Linear Response Regime

The slow-motional regime for a tumbling spin probe can be defined operationally as the range of molecular motional rates where a change in motional rate gives rise to an observable effect on the ESR spectra but the spectra cannot be adequately described by a fast-motional theory. The ESR spectra in the fast-motional limit are well understood but are less sensitive to the details of the motions of the spin probes than slow-motional spectra, whereas the rigid limit provides no motional information whatsoever. Thus, one is forced to tackle the problem of the interpretation of the slow-motional spectra.

The breakdown of the fast-motional theories can be traced to their perturbative nature. In the slow-motional regime the dynamics of the spins are strongly coupled to the orientational and/or positional degrees of freedom of the molecule, which render perturbative treatments invalid. To proceed, one must treat both the classical orientational and/or positional degrees of freedom and the quantum-mechanical spin degrees of freedom on a more equal footing. Since solving the exact equations of motion for all the molecules in the sample is obviously an impossible task, some physically reasonable assumptions must be introduced to make the problem tractable.

First, assume that the equation of motion for the density matrix, $\rho(t)$, has the same Hamiltonian $\hat{\mathscr{H}}(t)$ for all members of the ensemble and is given by the

quantum-mechanical Liouville equation

$$\frac{\partial \rho}{\partial t} = -i[\hat{\mathscr{H}}(t), \rho], \tag{4}$$

where $\hat{\mathscr{H}}(t)$ is given in angular frequency units.

Now, assume that the time dependence of the spin Hamiltonian $\hat{\mathscr{H}}(t)$ for a spin probe arises from interactions with its environment such that $\hat{\mathscr{H}}(t)$ is fully determined by a complete set of random variables, $\Omega$. Also assume that this time dependence of $\Omega$ is described by a stationary Markov process, so that the probability of being in a state $\Omega_2$ at time $t_2$, if in state $\Omega_1$ at time $t_1 = t_2 - \Delta t$, is independent of the value of $\Omega$ at any time earlier than $t_1$ and depends only on the time difference $\Delta t$ and not on $t_1$. A stationary Markov process can be described by a differential equation,

$$\frac{\partial P(\Omega, t)}{\partial t} = -\Gamma(\Omega) P(\Omega, t), \tag{5}$$

where $P(\Omega, t)$ is the probability of the spin probe being in a state $\Omega$ at time $t$.

Since the process is assumed stationary, $\Gamma(\Omega)$ is independent of time. The stochastic evolution operator $\Gamma(\Omega)$ operates only on the random variables $\Omega$ and is independent of the spin degrees of freedom and may include such general Markov operators as the diffusion operators given by Fokker–Planck equations and transition rate matrices among discrete states. In most of our examples, $\Omega$ will represent Euler angles specifying orientation and $\Gamma(\Omega)$ will be a rotational diffusion or Fokker–Planck operator. It is also assumed that the stochastic process has a unique equilibrium distribution $P_0(\Omega)$ characterized by

$$\Gamma(\Omega) P_0(\Omega) = 0. \tag{6}$$

It can be shown [5, 7, 50] that Eqs. 4–6 lead to the SLE of motion,

$$\begin{aligned} \frac{\partial \rho(\Omega, t)}{\partial t} &= -i[\hat{\mathscr{H}}(\Omega), \rho(\Omega, t)] - \Gamma(\Omega)\rho(\Omega, t) \\ &= -i\mathbf{L}(\Omega)\rho(\Omega, t) - \Gamma(\Omega)\rho(\Omega, t), \end{aligned} \tag{7}$$

where $\rho(\Omega, t)$ is now understood to be the value of the density matrix associated with a particular value of $\Omega$ and hence of $\hat{\mathscr{H}}(\Omega)$. Thus, instead of looking at the explicit time dependence of the spin Hamiltonian $\hat{\mathscr{H}}(t)$ involving the interaction with its environment, the spin Hamiltonian is written



in terms of random variables $\Omega$, and their modulation (e.g., due to rotational motions) is expressed by the time dependence of $\Omega$.

Equation 7 implicitly neglects the back reaction of the spins on the random variables $\Omega$, so spin relaxation induced by the coupling of the spins to $\Omega$ will tend to relax the spins to infinite temperature. This is not a concern for the lineshape problem (cf. Eq. 1). The solution is now well known and will be discussed further in Section VII, where pulsed and nonlinear phenomena are treated, and this matter is important.

The general linear response expression for the imaginary part of the magnetic susceptibility $\chi''(\omega)$ resulting from a very weak linearly polarized microwave field of angular frequency $\omega$ being applied to the system [51] is

$$\chi''_{jj}(\omega) = \frac{\omega}{2Nk_BT} \int_0^\infty dt (e^{i\omega t} + e^{-i\omega t}) \operatorname{Tr}\{\mathscr{M}_j(t)\mathscr{M}_j\}, \tag{8}$$

which involves a trace over the macroscopic magnetization operator $\mathscr{M}_j$. In Eq. 8, $T$ is the absolute temperature, $k_B$ Boltzmann's constant, and $N$ is the number of spin eigenstates of the spin probe. Note that $M$, the macroscopic value of the magnetization is related to $\mathscr{M}$, its associated quantum-mechanical operator, by $M(t) = \operatorname{Tr}\{\rho(t)\mathscr{M}\}$. The oscillating field is taken along the $j = x, y$, or $z$ direction. For our system of noninteracting (or weakly interacting) spin probes with nearly isotropic $g$ value, we have

$$\operatorname{Tr}\{\mathscr{M}_j(t)\mathscr{M}_j\} = \mathscr{N}\gamma_e^2 \operatorname{Tr}\{S_j(t)S_j\} = \mathscr{N}\gamma_e^2 \operatorname{Tr}\{S_jS_j(t)\}, \tag{9}$$

where $\mathscr{N}$ is the number of spins in the sample and $\gamma_e$ is the magnetogyric ratio of the electron. The spin operator $S_j(t)$ in the Heisenberg representation will obey a SLE equivalent to Eq. 7 given by

$$\frac{\partial S_j(\Omega, t)}{\partial t} = i\mathbf{L}(\Omega)S_j(\Omega, t) - \Gamma^\dagger(\Omega)\rho(\Omega, t), \tag{10}$$

where the superscript dagger implies the Hermitian adjoint operator. This is required to have the expectation values of the magnetization be identical in the Heisenberg and Schrödinger pictures given by Eqs. 10 and 7, respectively. The SLE in Eq. 10 is subject to the initial condition

$$S_j(\Omega, 0) = P_0(\Omega)S_j. \tag{11}$$

This form will be needed in order to interpret Eq. 9 as a proper equilibrium-averaged correlation function.

Now in Eqs. 8 and 9, the trace over orientational degrees of freedom is

replaced by a classical average indicated by an overbar:

$$\overline{S_j S_j(t)} = \overline{S_j S_j(\Omega, t)} = \int d\Omega\, S_j S_j(\Omega, t). \tag{12}$$

In this notation

$$\chi''_{jj}(\pm\omega) = \frac{\mathcal{N}\gamma_e^2\omega}{2Nk_BT}\,\mathrm{Tr}_s\{\overline{S_j S_j(\pm\omega, \Omega)}\}, \tag{13}$$

where the trace is only over spin degrees of freedom, $S_j(\pm\omega, \Omega)$ is the Fourier–Laplace transform of $S_j(\Omega, t)$,

$$S_j(\pm\omega, \Omega) \equiv \int_0^\infty dt\, e^{\mp i\omega t} S_j(\Omega, t), \tag{14}$$

and the plus and minus signs are found to correspond to the two counterrotating components of the microwave field, only one of which is important in large static magnetic fields.

From Eqs. 10 and 14 it follows that

$$S_j(\omega, \Omega) = [i(\omega\mathbf{I} - \mathbf{L}) + \mathbf{\Gamma}^\dagger]^{-1} S_j(0, \Omega). \tag{15}$$

Thus,

$$\overline{S_j S_j(\pm\omega, \Omega)} = \int d\Omega\, S_j [i(\omega\mathbf{I} - \mathbf{L} + \mathbf{\Gamma}^\dagger]^{-1} P_0(\Omega) S_j. \tag{16}$$

This expression may be inserted in Eq. 13 to obtain $\chi''_{jj}(\pm\omega)$.

It is convenient at this point to introduce a "symmetrizing" transformation for the stochastic Liouville operator. It is not needed for isotropic liquids with simple models but becomes useful for anisotropic liquids [9] or more sophisticated models [15]. The relevant similarity transformation is

$$\tilde{\mathbf{\Gamma}}(\Omega) = P_0^{-1/2}(\Omega)\mathbf{\Gamma}(\Omega)P_0^{1/2}(\Omega), \tag{17}$$

where here $P_0^{\pm 1/2}(\Omega)$ are regarded as operators. This transformation defines $\tilde{\mathbf{\Gamma}}$ in a form that may be represented by a symmetric matrix that is, in general, complex (cf. Sections VI.B and VI.B.1). The symmetrized diffusion operator $\tilde{\mathbf{\Gamma}}(\Omega)$ will be used in the remainder of this section.

Equation 16 may be rewritten in more symmetric form as

$$\overline{S_j S_j(\pm\omega, \Omega)} = \int d\Omega\, S_j P_0^{1/2}(\Omega) [i(\omega\mathbf{I} - \mathbf{L} + \tilde{\mathbf{\Gamma}}^\dagger]^{-1} P_0^{1/2}(\Omega) S_j. \tag{18}$$



The trace over spin space and averaging over the ensemble in Eq. 18 may be regarded as a scalar product in Liouville space that can be represented as

$$\mathrm{Tr}_s\overline{S_j S_j(\pm\omega, \Omega)} = \langle P_0^{1/2} S_j | [i(\omega\mathbf{I} - \mathbf{L}) + \tilde{\mathbf{\Gamma}}^\dagger]^{-1} | P_0^{1/2} S_j \rangle, \tag{19}$$

which has the form of Eq. 1.

### B. The Spin Hamiltonian

The total spin Hamiltonian $\hat{\mathscr{H}}(t)$, expressed in angular frequency units, can be separated into three components,

$$\hat{\mathscr{H}}(t) = \hat{\mathscr{H}}_0 + \hat{\mathscr{H}}_1(\Omega) + \hat{\mathscr{E}}(t). \tag{20}$$

In the high-field approximation the orientation-independent component $\hat{\mathscr{H}}_0$,

$$\hbar\hat{\mathscr{H}}_0 = \gamma_e B_0 \hat{S}_z - \hbar\sum_i \gamma_i B_0 \hat{I}_z + \hbar\gamma_e \sum_i a_i \hat{S}_z \hat{I}_z, \tag{21}$$

gives the zero-order energy levels and transition frequencies. The orientation-dependent part, $\hat{\mathscr{H}}_1(\Omega)$, can be expressed as the scalar product of two tensors [52]:

$$\hat{\mathscr{H}}_1(\Omega) = \sum_{\mu,i}\sum_{L,M,K} (-1)^K F'^{(L,-K)}_{\mu,i} \mathscr{D}^L_{KM}(\Omega) A^{(L,M)}_{\mu,i}, \tag{22}$$

where the $F'^{(L,K)}_{\mu,i}$ and $A^{(L,M)}_{\mu,i}$ are irreducible tensor components of rank $L$. The $F'^{(L,K)}_{\mu,i}$ are spatial functions in molecule-fixed coordinates, whereas $A^{(L,M)}_{\mu,i}$ is a spin operator defined in the laboratory axis system. The subscripts $\mu$ and $i$ refer to the type of perturbation and to the different nuclei, respectively. The generalized spherical harmonics $\mathscr{D}^L_{KM}(\Omega)$ include the transformation from the molecule-fixed axis system $(x', y', z')$ into the laboratory axis system $(x, y, z)$. For the analysis of most slow-motional ESR spectra of simple free radicals where $S = 1/2$, only second-rank tensors are important, for example, the $A$- and $g$-tensors. The calculation of simple types of matrix elements of the Liouville operator derived from $\hat{\mathscr{H}}_0 + \hat{\mathscr{H}}_1$ is summarized in Appendix D.

### C. Model Diffusion Operators

When the general method is applied to rotational modulation, $\Omega$ can be taken to be the Euler angles for a molecular axis system fixed to a tumbling spin probe with respect to a fixed laboratory axis system. For a molecule undergoing many collisions, causing small random angular reorientations, the resulting isotropic Brownian rotational motion is a Markov process, which

can be described by the rotational diffusion equation

$$\frac{\partial P(\Omega, t)}{\partial t} = -R\nabla_\Omega^2 P(\Omega, t), \tag{23}$$

where $\nabla_\Omega^2$ is the Laplacian operator on the surface of the unit sphere and $R$ is the rotational diffusion coefficient.

In an isotropic liquid, the equilibrium probability $P_0(\Omega)$ given by Eq. 6 will be equal for all orientations, so that $P_0(\Omega) = 1/8\pi^2$. Here the Markov operator $\Gamma(\Omega)$ for isotropic Brownian rotation is independent of $\Omega$ since the liquid is assumed to be isotropic. The operator $\Gamma = R\nabla_\Omega^2$ is of the form of the Hamiltonian for a spherical top; therefore, its orthonormal eigenfunctions are the generalized spherical harmonics

$$\phi_{LMK}(\Omega) = \left(\frac{2L+1}{8\pi^2}\right)^{1/2} \mathcal{D}^L_{MK}(\Omega) \tag{24}$$

with eigenvalues $RL(L+1)$ [53, 54].

Similarly, the Markov operator for axially symmetric Brownian diffusion about a molecule-fixed $z$ axis is formally the Hamiltonian for a symmetric top whose symmetry axis is the $z$ axis. The orthonormal eigenfunctions are again the normalized Wigner rotation matrices with eigenvalues $R_\perp L(L+1) + (R_\parallel - R_\perp)K^2$, where $R_\perp$ and $R_\parallel$ are the rotational diffusion constants about the $x$ and $y$ and about the $z$ axes, respectively [7, 12, 53, 54]. The "quantum numbers" $K$ and $M$ of the Wigner rotation matrices refer to projections along the body-fixed symmetry axis and along a space-fixed axis, respectively. For completely asymmetric Brownian rotation the diffusion constants about the three principal axes are all unequal, and the stochastic operator has more complicated solutions [7, 12, 53, 54].

Some canonical models for rotational reorientation frequently used in ESR spectroscopy are the following

- Brownian rotational diffusion.
- An approximation to free diffusion in which a molecule rotates freely for time $\tau$ (i.e., inertial motion with $\tau = I/\beta$, where $I$ is the moment of inertia and $\beta$ is the friction coefficient) and then reorients instantaneously.
- Jump diffusion in which a molecule has a fixed orientation for a fixed time $\tau$ and then "jumps" instantaneously to a new orientation with no inertial effects [13, 55].

For isotropic reorientation, the characteristic relaxation rates (eigenvalues) associated with the generalized spherical harmonic eigenfunctions of rank $L$ are all degenerate for these three models and are given, respectively, by

$$\tau_L^{-1} = \frac{RL(L+1)}{[1 + R\tau L(L+1)]^{1/2}},$$

$$\tau_L^{-1} = \tau^{-1}\left[1 - \frac{1}{2L+1}\int_0^\pi d\psi\, W(\psi)\frac{\sin(L+1/2)\psi}{\sin(\psi/2)}\right],$$

where $W(\psi)$ is the distribution function for diffusive steps by an angle $\psi$ about an arbitrary axis and is normalized so that

$$\int_0^\pi d\psi\, W(\psi) = 1.$$

More sophisticated Markovian models and the general matter of modeling the molecular dynamics are discussed in Section VI. Some of the simple types of matrix elements of $\tilde{\Gamma}$ that arise in anisotropic liquids are given in Appendix D.

## III. OVERVIEW OF DISCRETIZATION METHODS FOR THE SOLUTION OF THE SLE

For typical forms of the diffusion operator, the SLE is a set of coupled partial differential equations (PDEs) governing the time evolution of the orientation-dependent quantum-mechanical spin density matrix subject to specific initial conditions. This set of PDEs can be simplified to a set of coupled linear algebraic equations (LAEs) by Fourier–Laplace transformation of the set of PDEs with respect to time (cf. Eq. 1) followed by discretization of the spatial parts. The discretization is necessary to remove the spatial derivative terms usually present in the diffusion operator. The resulting set of LAEs can then be solved in a variety of ways. The proper choice of method of solution depends on the structure of the matrix and the specific quantity desired.

There are three important techniques to achieve discretization of angular-dependent terms: expansion in a set of basis functions, finite-difference approximation, and finite-element approximation. All of these methods typically give rise to very large sparse matrices characterizing the stochastic Liouville operator **A**.

The expansion in a set of global basis functions is analogous to the method of variation of constants used in elementary quantum mechanics and the study of differential equations. The set of basis functions used in the expansion is usually not complete in the formal sense since this would imply an infinite set

of coupled equations, but it can be made adequate for numerical calculations to any degree of accuracy desired. The most common set of functions to use as basis functions in cases involving rotational diffusion are the generalized spherical harmonics. There are several reasons for this choice:

- They are the eigenfunctions of the three canonical rotational Fokker–Planck operators in isotropic liquids (cf. Section II).
- They, by definition, have well-defined transformation properties with respect to coordinate rotations; thus, powerful group-theoretic techniques can be used to simplify calculations.
- Their eigenvalues have favorable scaling properties with respect to the principal quantum number $L$.
- They form an orthonormal set, and the infinite set of all generalized spherical harmonics form a basis in which any square integrable function on the unit sphere in four space can be expanded.

Though these functions are the eigenfunctions of the quantum-mechanical rigid rotator Hamiltonian, it is important to realize that these functions are only used to expand the orientation-dependent density matrix in this application and have nothing to do with angular momentum! Nevertheless, all the sophisticated angular momentum coupling and transformation techniques that have been developed in other areas of physics and chemistry can be applied since these functions have well-defined properties under rotations. The properties of the generalized spherical harmonics and applications of angular momentum techniques to density matrix problems are described in detail in various texts [56, 57] and papers [58–64].

For example, the operators $S_j(\Omega, \omega)$ given in Eq. 15 can be expanded as

$$|S_j(\Omega, \omega)\rangle = \sum_{\lambda,m} s_{j,\lambda,m}(\omega)|\lambda, m\rangle, \tag{25}$$

where the ket $|\lambda, m\rangle = |\lambda\rangle|m\rangle$ is a product of spin operators and spatial functions, where $\lambda$ represents labels for the spin operator basis, $m$ represents the labels associated with the basis of spatial functions, and the expansion coefficients $s_{j,\lambda,m}(\omega)$ are, in general, complex-valued functions. In this notation Eq. 19 becomes

$$\mathrm{Tr}_s\overline{S_j S_j(\pm\omega, \Omega)} \propto \mathbf{v}^{\mathrm{tr}}[\mathbf{A}'(\omega)]^{-1}\mathbf{v}, \tag{26}$$

where $\mathbf{A}'(\omega)$ is the matrix of the operator $i(\omega\mathbf{I} - \mathbf{L}) + \tilde{\mathbf{\Gamma}}^{\dagger}$ in the basis $|\lambda, m\rangle$, the superscript tr means transpose, and the elements of the starting vector are

$$v_{\lambda,m} = \langle\lambda, m|P_0^{1/2}S_j\rangle = \frac{1}{N}\mathrm{Tr}\{S_\lambda^\dagger S_j\}\langle m|P_0^{1/2}(\Omega)\rangle. \tag{27}$$



Equation 26 is equivalent to Eq. 1. This discretization procedure can be used to generate a complex symmetric form for $\mathbf{A}'(\omega)$. The evaluation of the components of $\mathbf{v}$ is discussed in Section V.A and in Appendix D.

The finite-difference approximation involves explicit discretization of the spatial variables. By assuming that the radical can only be found at these discrete positions or orientations, it is possible to approximate the Fokker–Planck equation for the particle by a finite-difference equation. This is a very popular technique for solving partial differential equations in many areas of science. The finite-difference approach to solving the SLE was used by Gordon and Messenger for angular variables [45] and by Freed and co-workers for translational problems [65–70], and is still in widespread use [71–76].

The finite-element approach involves approximating the solution of the SLE in a piecewise fashion over finite areas on the unit sphere [77–79] or over finite volumes in Cartesian space [78, 79]. Usually, the solution is well approximated by low-order polynomial functions, and appropriate continuity requirements are enforced along the boundaries of the elements. This matching of elements at the boundaries implies that these functions do not form an orthogonal set, but the solution is not uniquely defined without these conditions. The lack of orthogonality means that the computer solution of the equations generated by applying the finite-element approximation is more difficult by traditional means, as it leads to a generalized eigenvalue problem [80, 81]. The use of finite elements for the SLE and the associated variational problem are discussed in detail by Zientara and Freed [78]. The alternative approach of using global, orthogonal functions instead of piecewise smooth polynomials over small regions is identical to the basis function expansion method already discussed. Derived in this manner, it is known as the global Galërkin variational method [78, 81].

The most prominent exceptions to the pattern of discretization followed by matrix manipulation are the Monte Carlo methods developed by Pedersen [82] and Itzkowitz [83] where the relaxation function is evaluated directly and the spectrum is obtained by Fourier transformation. Though this method seems to be less efficient on conventional computers than the matrix-oriented approaches discussed, the popularity of the Monte Carlo technique in other disciplines has spurred the development of new computer architectures and associated algorithms that should prompt renewed interest in this approach for the CW lineshape problem. The remaining drawback to the Monte Carlo approach, namely, the inaccessibility of the eigenvalue–eigenvector decomposition, makes it inapplicable for spin echo calculations and other applications where these quantities are required.

It is also possible to use numerical integration techniques on the discretized equations of motion to directly evaluate the time evolution, but this has only been attempted when the complete time dependence is required, such as in the

detailed investigation of the interaction of the spins with microwave pulses of finite amplitude and duration [84] or when the spin Hamiltonian has a particularly simple form [85].

## IV. OVERVIEW OF TRADITIONAL MATRIX METHODS FOR THE SOLUTION OF THE SLE

The algorithms commonly used for the calculation of slow-motional spectra from the SLE fall into two main categories based on efficiency and ability to handle large sparse matrices. First are the traditional algorithms for diagonalizing matrices and solving sets of coupled linear equations. Second, there are the various forms of the LA for tridiagonalization and the related conjugate gradients algorithm for solving sets of linear equations.

The first class of traditional methods is characterized by a variety of difficulties and strengths. These algorithms typically amount to "computational overkill" for the problem at hand. For instance, the Rutishauser–$QR$ diagonalization gives the full set of eigenvectors and eigenvalues though only a small subset are important in the final spectrum. This large computational overhead and their characteristic of modifying the sparsity structure of $\mathbf{A}$ combine to make these algorithms unattractive for present purposes. An important strength of this class of algorithms is their well-characterized stability and reliability.

In contrast, the LA and its kin are much better suited for the efficient calculation of magnetic resonance spectra. They are effective in handling large sparse matrices since they do not modify the original matrix. It is therefore possible to take advantage of the very special sparsity structure of the stochastic Liouville matrix. In addition, all of the quantities calculated in the LA are either used directly in the calculation of the spectrum or are needed in the next recursive step (see Section V.A and Appendix A). In this sense, the LA represents a good approximation to the minimal amount of computation necessary to compute magnetic resonance spectra in the linear response regime.

Before getting into a discussion on the computational aspects of solving the SLE by matrix methods, it is valuable to review the analytic aspects of the problem. In general, the stochastic Liouville matrix $\mathbf{A}$ can always be represented as a complex symmetric matrix (CSM) [3], and $\mathbf{A}$ *cannot* be Hermitian in the presence of relaxation. In addition, there is a band outside of which all matrix elements are identically zero. The class of CSM is quite general, and many of the theorems on diagonalization and related topics do not have simple analogues for non-Hermitian CSMs [86, 87]. For instance, one is not guaranteed that an arbitrary CSM can be diagonalized by a similarity transformation. However, if a CSM is diagonalizable by a similarity transformation, it is diagonalized by a complex orthogonal matrix (COM) $\mathbf{O}$:

$$\mathbf{O}^{tr}\mathbf{A}\mathbf{O} = \mathbf{\Lambda}, \tag{28}$$

where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues $\lambda_i$ of $\mathbf{A}$. We will *assume* that the stochastic Liouville matrices under consideration are diagonalizable. This is not too drastic an assumption in light of the fact that any square matrix is arbitrarily close to a diagonalizable matrix [87]. The pathologies that can result from nondiagonalizability have been thoroughly studied [86, 87] and only introduce irrelevant complications into the present discussion. The class of COM is also peculiar in many respects. In particular, the magnitude of the elements of a COM is not bounded as is the case for unitary matrices familiar from quantum mechanics. Since $A$ is non-Hermitian, its eigenvalues are not constrained to lie on the real axis. In spite of the lack of the simple behavior characteristic of Hermitian matrices, we can say something useful about the regions in the complex plane that can and cannot contain eigenvalues. Since the diffusion operator is nonnegative (i.e., $\langle\phi|\mathbf{\Gamma}|\phi\rangle \geq 0$), a well-known theorem from linear algebra states that all the eigenvalues of $\mathbf{A}$ must lie in the closed right half of the complex $\lambda$ plane [87]. In physical terms, this corresponds to the fact that the relaxation must force the system toward equilibrium. It is very important to note that the localization of eigenvalues to the right-half plane is independent of the dimension of the basis set (i.e., it does not depend on the number of basis functions, finite differences, or finite elements used in the discretization). In addition, the eigenvalues are restricted, by a similar argument, to lie within a band about the real axis, since the spectrum of the Liouville operator is bounded. The width of the band is dependent on the rigid-limit magnetic tensors. Other eigenvalue localization theorems such as Gerschgorin's theorem can be applied, but they do not give rise to transparently useful results except in the fast-motional limit [88], though they do form a basis for some of the traditional diagonalization algorithms. References to several good treatments of the properties of complex orthogonal matrices can be found in the bibliography [86, 87, 89–94].

We will briefly survey the traditional methods for diagonalization and the solution of linear systems of equations from the point of view of calculating magnetic resonance spectra. This is not intended to be a comprehensive or definitive treatment of these topics. For much more information about the theory and usage of the algorithms discussed here, the reader is encouraged to consult standard texts on numerical linear algebra [80, 95, 96] and the references therein.

### A. Solving Linear Systems of Equations Using Gaussian Elimination

The most useful classical method of solving the type of linear systems of equations that arise in the calculation of magnetic resonance spectra is

Gaussian elimination with partial pivoting. This method has been used in the past for the calculation of CWESR spectra [13,97] and is still used in the quantitative analysis of saturation transfer spectra [98,99] and the calculation of spectral densities in paramagnetic NMR [100].

The idea behind the Gaussian elimination method is to factor the matrix $\mathbf{A}'(\Delta\omega)$ in Eq. 3 as a product of a lower triangular matrix $\mathbf{L}$ and an upper triangular matrix $\mathbf{U}$, with care taken to arrange the sequence of operations to minimize loss of accuracy from the finite-precision computer arithmetic. In this manner, one solves Eq. 3 by successively solving two simpler triangular systems and never generating the inverse matrix explicitly.

The procedure can be summarized by rewriting Eq. 3 in the form

$$\mathbf{PA}'(\Delta\omega)\mathbf{u}(\Delta\omega) = \mathbf{LU}\mathbf{u}(\Delta\omega) = \mathbf{v}, \tag{29}$$

where $\mathbf{P}$ is a permutation matrix that arises from the sequencing of the operations. The value of factorizing the matrix lies in the fact that the solution of Eq. 3 can be broken down into solving the equation

$$\mathbf{Ly} = \mathbf{P}^{tr}\mathbf{v} \tag{30}$$

for $\mathbf{y}$ followed by solving

$$\mathbf{U}\mathbf{u}(\Delta\omega) = \mathbf{y} \tag{31}$$

for the desired vector $\mathbf{u}(\Delta\omega)$. Equations 30 and 31 involve triangular matrices and hence can be solved in $\mathcal{O}(N^2)$ floating-point operations. The $\mathbf{LU}$ factorization itself is, however, an $\mathcal{O}(N^3)$ process.

The Gaussian elimination method is quite stable and reliable if the matrix $\mathbf{A}'(\Delta\omega)$ is neither singular nor nearly so. It has the definite disadvantage that it requires large amounts of computer time [$\mathcal{O}(N^3)$ operations] and memory (all elements within the bandwidth must be stored). The complete procedure must be carried out for each value of the frequency at which the value $I(\Delta\omega)$ is desired, since the $\mathbf{LU}$ factorization is not independent of $\Delta\omega$.

### B. Complete Diagonalization Methods

There are two main algorithms for the eigenvalue–eigenvector decomposition of the type of complex symmetric matrices that arise in magnetic resonance problems. The first is a variant of an algorithm devised by Jacobi for the direct diagonalization of real symmetric matrices. The second, more efficient algorithm is a variant of Given's method [101], due to Rutishauser [45, 102], which eliminates some of the drawbacks of the Jacobi algorithm by first tridiagonalizing the matrix and then using the $QR$ iteration to diagonalize the tridiagonal matrix.



The complete diagonalization of a given matrix is usually much more time consuming than solving Eq. 3 by Gaussian elimination with partial pivoting for a single value of $\Delta\omega$. Therefore, it makes sense to use a diagonalization method if $\mathbf{A}'(\Delta\omega) = \mathbf{A} - i\,\Delta\omega\,\mathbf{I}$, since in this case the spectrum can be easily computed for any value of $\Delta\omega$. The formula for $I(\Delta\omega)$ can be easily derived by using the diagonalizing transformation (see Eq. 28) to rewrite Eq. 1 as

$$\begin{aligned} I(\Delta\omega) &= \frac{1}{\pi}\,\mathrm{Re}\left\{\mathbf{v}^{tr}\mathbf{O}\mathbf{O}^{tr}[\mathbf{A} + i\,\Delta\omega\,\mathbf{I}]^{-1}\mathbf{O}\mathbf{O}^{tr}\mathbf{v}\right\} \\ &= \frac{1}{\pi}\,\mathrm{Re}\left\{(\mathbf{O}^{tr}\mathbf{v})^{tr}[\mathbf{A} + i\,\Delta\omega\,\mathbf{I}]^{-1}(\mathbf{O}^{tr}\mathbf{v})\right\}. \end{aligned} \tag{32}$$

Since $\mathbf{\Lambda} + i\,\Delta\omega\,\mathbf{I}$ is diagonal, this can be collapsed to a simple sum over the eigenvalues of the form

$$I(\Delta\omega) = \frac{1}{\pi}\,\mathrm{Re}\left\{\sum_{j=1}^{N}\frac{c_j^2}{\lambda_j + i\,\Delta\omega}\right\}, \tag{33}$$

where $c_j = (\mathbf{O}^{tr}\mathbf{v})_j$ is the projection of the $j$th eigenvector on the starting vector. The LAEs derived by finite difference or expansion in a set of trial functions typically have this form, though the finite-element method usually gives rise to off-diagonal elements containing $\omega$ because of the nonorthogonality of the piecewise polynomial basis functions.

#### 1. *Jacobi-Type Methods*

The Jacobi algorithm involves the successive zeroing of the largest off-diagonal matrix element by (real) orthogonal similarity transformations. Unfortunately, these successive transformations tend to fill in matrix elements that were initially zero; thus, this algorithm is not too well suited for sparse matrix diagonalization. In spite of the proliferation of nonzero matrix elements, it can be shown that the rotations can be chosen such that the sum of the moduli of the off-diagonal matrix elements is reduced at every stage. The product of these orthogonal matrices is the matrix of eigenvectors while the reduction in the sum of the moduli of the off-diagonal matrix elements causes the original matrix to converge to the desired diagonal matrix of eigenvalues by Gerschgorin's theorem.

The original algorithm, which calls for the selective zeroing of the largest off-diagonal element, requires a complete search of the off-diagonal matrix elements at every stage. This searching is very costly in terms of computer time and is not too productive. In practice, the so-called cyclic Jacobi algorithm is preferred since it does not require searching the off-diagonal matrix elements.

In the cyclic Jacobi algorithm one starts at the diagonal and moves across a row, zeroing each matrix element in turn. When the end of the row is reached, the same procedure is applied to the next row. When the end of the matrix is reached, the entire procedure is repeated until all of the off-diagonal matrix elements are below some small value characteristic of the roundoff error of the computer.

Several authors have discussed the generalization of the cyclic Jacobi algorithm to handle complex symmetric matrices [103–106]. This generalization is rather straightforward.

All variants of the Jacobi algorithm suffer from two major flaws when applied to the large sparse matrices commonplace in magnetic resonance calculations. First, as mentioned previously, matrix elements that were originally zero get replaced by nonzero entries as the algorithm proceeds; thus, these algorithms have storage requirements that scale as the square of the dimension of the matrix to be diagonalized. Second is the fact that the number of Jacobi rotations needed to diagonalize a matrix can be infinite! Because of these drawbacks, the Jacobi methods have been superceded by alternative methods that rely on a reduction to tridiagonal form as an intermediate stage. Nevertheless, there are applications where the Jacobi algorithm is quite useful [80].

### 2. *Tridiagonalization and the QR Algorithm*

In most applications it is faster to diagonalize a matrix by first reducing it to tridiagonal form followed by the diagonalization of the resulting tridiagonal matrix than it is to reduce it to diagonal form directly via the Jacobi method. The tridiagonalization method developed by Givens [101] also uses Jacobi rotations to zero off-diagonal matrix elements. The difference between the Jacobi diagonalization and the Givens tridiagonalization algorithms is the sequence in which the off-diagonal matrix elements are annihilated. It can be shown that the Givens tridiagonalization of a symmetric matrix can be accomplished by a finite number of rotations. A variant of the Givens method due to Rutishauser [102] is particularly well suited to the task at hand since it takes advantage of the banded nature of the stochastic Liouville matrix, thereby reducing storage requirements. This tridiagonalization method was popularized in the magnetic resonance community by the work of Gordon and Messenger [45].

The symmetric tridiagonal matrix generated by the Rutishauser algorithm can be diagonalized in several ways. The most common choice, also popularized by Gordon and Messenger [45], is the $QR$ algorithm. Details of the symmetric tridiagonal $QR$ algorithm can be found in standard references on numerical linear algebra [80, 96]. Recently, Cullum and Willoughby [89] have advocated an alternative procedure based on a $QL$ decomposition.



Prior to the introduction of Lanczos-based methods by Moro and Freed [3], the Rutishauser tridiagonalization–$QR$ iteration method was the most efficient method of calculating magnetic resonance spectra. It is, however, a "brute-force" approach in the sense that a large portion of the effort is expended calculating irrelevant eigenvalues whose corresponding eigenvectors have negligible overlap with the starting vector (see Eq. 33). There is no way for the algorithm to differentiate between important and unimportant eigenvalues as the computation proceeds. All eigenvalues and eigenvectors are on an equal footing, and the number of eigenvalues calculated must equal the dimension of the matrix. In contrast, this ability to differentiate between important and unimportant vectors is inherent in the Lanczos-based methods discussed in the next section.

## V. LANCZOS AND CONJUGATE GRADIENTS METHODS OF SOLVING THE SLE

The Lanczos algorithm (LA) and the related conjugate gradients algorithm are extremely effective for calculating slow-motional magnetic resonance spectra [3, 4, 20, 71]. In this section we will discuss the basic LA as applied to Hermitian and complex symmetric matrices. More sophisticated variations of the LA involving selective reorthogonalization [80, 96] and the identification of spurious and duplicated eigenvalues [89] have been developed to circumvent known numerical instabilities. The basic algorithms are sufficient for most magnetic resonance calculations.

There are several advantages to employing the LA for sparse matrix problems such as those presented by Eqs. 1 and 3. These advantages include the following:

- Only the nonzero matrix elements need to be stored; although, in general, they may be recomputed as needed to minimize computer memory requirements.
- The original matrix elements are not altered during the operations.
- No new matrix elements are created.
- The entire algorithm can easily be written in less than 100 lines of FORTRAN code.
- The recursive steps or projections on which the algorithm is based are very closely related to the projection methods in statistical mechanics.

These very practical considerations alone are sufficient reason to prefer the Lanczos methods to Gaussian elimination or complete diagonalization methods.

The theoretical advantages of the LA are as important as they are practical

to the overall success of the method. The LA provides a convenient conceptual framework for the identification and classification of the important physical features of the lineshape calculation problem through its relationship with the powerful memory function methods of nonequilibrium statistical mechanics [3, 107–112]. The traditional methods for calculating magnetic resonance spectra are based on theorems from linear algebra rather than physical insight. They are, by their very nature, nearly devoid of insight other than that contained in the specification of the terms in the stochastic Liouville matrix and the discretization method.

### A. The Lanczos Algorithm

The LA proceeds by recursive projections or steps that produce successively larger tridiagonal matrix approximations to the original matrix. These projections define the so-called Lanczos vectors. If $N$ is the dimension of the matrix and $n_s$ the number of recursive steps needed to converge to an accurate spectrum, then in all cases studied to date $n_s \ll N$. 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 $N$ that is spanned by the starting basis set of orthonormal vectors $|f_j\rangle$, $j = 1, 2, \ldots, N$, a smaller subspace spanned by the Lanczos vectors (i.e., the basis vectors for the tridiagonalized form of $\mathbf{A}$, which is $\mathbf{T}_n$) written as $|\Phi_k\rangle$, $k = 1, 2, \ldots, n$. When $n = n_s$, these Lanczos vectors are a sufficient basis for accurately representing the spectrum. In this sense the LA constructs subspaces that progressively approximate the "optimal reduced space" for the problem. These subspaces, spanned by the Lanczos vectors, are the Krylov subspaces [80, 89] generated by the span of the vectors $\mathbf{A}^{k-1}|v\rangle$ for $k = 1, 2, \ldots, n$. Thus, the choice of $|v\rangle$ as the starting vector for the LA 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 projection on $|v\rangle$ in exact arithmetic.

Now, consider the recursive steps of the LA. First, identify the starting vector $|v\rangle$ as the first Lanczos vector $|\Phi_1\rangle$ in accordance with the preceding discussion. A Gram–Schmidt orthogonalization on the Krylov sequence $\mathbf{A}^{k-1}|v\rangle$ for $k = 1, 2, \ldots, n$ recursively generates the set of orthornormal Lanczos vectors $|\Phi_k\rangle$ defined as

$$\beta_k|\Phi_{k+1}\rangle = (\mathbf{I} - \mathbf{P}_k)\mathbf{A}|\Phi_k\rangle, \tag{34}$$

where $\beta_k$ is the normalizing coefficient chosen such that

$$\langle\Phi_{k+1}|\Phi_{k+1}\rangle = 1 \tag{35}$$

and $\mathbf{P}_k$ is the projection operator on the Krylov subspace spanned by the previous Lanczos vectors

$$\mathbf{P}_k = \sum_{j=1}^{k} |\Phi_j\rangle\langle\Phi_j| \qquad k \leqslant n. \tag{36}$$

Equation 34 leads to a three-term recursive relation for generating the $|\Phi_j\rangle$ (cf. Appendix A):

$$\beta_k|\Phi_{k+1}\rangle = (\mathbf{A} - \alpha_k\mathbf{I})|\Phi_k\rangle - \beta_{k-1}|\Phi_{k-1}\rangle, \tag{37}$$

where

$$\alpha_k = \langle\Phi_k|\mathbf{A}|\Phi_k\rangle \tag{38}$$

and

$$\beta_{k-1} = \langle\Phi_k|\mathbf{A}|\Phi_{k-1}\rangle. \tag{39}$$

It may easily be shown that $\mathbf{A}$ has an $n \times n$ tridiagonal approximation $\mathbf{T}_n$ in the basis of Lanczos vectors:

$$\langle\Phi_k|\mathbf{A}|\Phi_j\rangle = 0 \tag{40}$$

if $k \neq j, j \pm 1$, while Eqs. 38 and 39 give the nonzero matrix elements. That is, given the vectors $|\Phi_j\rangle$ in terms of their components $q_{j,k}$ in the original basis set $|f_j\rangle$, $j = 1, 2, \ldots, N$,

$$|\Phi_k\rangle = \sum_{j=1}^{N} q_{j,k}|f_j\rangle, \tag{41}$$

$$q_{j,k} = \langle f_j|\Phi_k\rangle, \tag{42}$$

the column vectors $\mathbf{q}_k$ form the matrix $\mathbf{Q}_n$ with orthonormal columns such that $\mathbf{Q}_n^{tr}\mathbf{Q}_n = \mathbf{I}_n$ and

$$\mathbf{T}_n = \mathbf{Q}_n^{tr}\mathbf{A}\mathbf{Q}_n. \tag{43}$$

This is the conventional single-vector LA for real symmetric matrices. The substitution of Hermitian conjugation for transposition in the preceding equations gives the analogous scheme for general Hermitian matrices.

For applications such as the calculation of magnetic resonance spectra and spectral densities associated with Fokker–Planck equations the matrix $\mathbf{A}$ is either complex symmetric or can be transformed to complex symmetric form [3, 44, 43] (cf. Section VI.D). Moro and Freed [3, 43] have shown that one can

simplify problems of this type by introducing a generalized norm and scalar product. That is, first consider the general non-Hermitian case. One can introduce a biorthonormal set of functions $|\Phi_j\rangle$ and $|\Phi^{j'}\rangle$ such that

$$\langle \Phi^{j'} | \Phi_j \rangle = \delta_{j',j}, \tag{44}$$

or, alternatively, letting $\mathbf{x}_j$ and $\mathbf{x}^{j'}$ be their column vector representations,

$$(\mathbf{x}^{j'})^{\dagger} \mathbf{x}_j = \delta_{j',j}. \tag{45}$$

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

$$\mathbf{x}^j = \mathbf{x}_j^* \tag{46}$$

such that Eq. 45 becomes

$$\mathbf{x}_{j'}^{tr} \mathbf{x}_j = \delta_{j',j}. \tag{47}$$

The Lanczos recursion method remains applicable with Eq. 47 defining the generalized scalar product. Note that the left vector is lacking the usual complex conjugation. The important aspects of this generalized scalar product are developed in Appendices A and C.

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 $\mathbf{A}$ [3]. This is obviously superior to the traditional methods that require $\mathcal{O}(N^3)$ time since $n_E$, $n_s \ll N$.

Finally, note that the tridiagonal form of the complex symmetric matrix $\mathbf{T}_n = \mathbf{Q}_n^{tr} \mathbf{A} \mathbf{Q}_n$ allows the application of very efficient diagonalization methods [45, 89]. The spectrum defined by Eq. 1 can easily be computed from using the eigenvalues of $\mathbf{T}_n$ and the projections of the associated eigenvectors on the starting vector. However, for computing CW spectra, a continued-fraction method [3, 44] can be used directly on the elements of the matrix $\mathbf{T}_n$. That is, since $|v\rangle$ is the first Lanczos vector and the Lanczos vectors are orthogonal in the sense of Eq. 47, the spectrum is given by

$$I(\Delta\omega) = \frac{1}{\pi} [i\,\Delta\omega\, \mathbf{I}_n + \mathbf{T}_n]_{1,1}^{-1}. \tag{48}$$

By examining the structure of the (1, 1) element of the inverse of successively larger principal submatrices of $[i\,\Delta\omega\, \mathbf{I}_n + \mathbf{T}_n]$, it is easy to show that the $I(\Delta\omega)$ can be written in the continued-fraction form (see Appendix B),

$$I(\Delta\omega) = \frac{1}{\pi} \mathrm{Re}\left[ \left( i\,\Delta\omega + \alpha_1 - \frac{\beta_1^2}{i\,\Delta\omega + \alpha_2 - \beta_2^2 \cdots} \right)^{-1} \right]. \tag{49}$$



The application of the LA to $\mathbf{A}$ generates the continued-fraction representation of the spectrum or spectral density. The same result, apart from the identification of $\tilde{\Gamma}$ with the classical Liouville operator, has been derived by H. Mori in the context of the dynamics of systems of interacting particles [107]. As a matter of fact, the same methodology, more specifically the recursive structure of Eq. 34, is the foundation of both the LA and Mori's derivation.

The relation specified by the continued fraction in Eq. 49 is quite general. Analytical calculation of the coefficients $\alpha_j$ and $\beta_j$ from the explicit operator form of $\mathbf{A}$ is possible and has been carried out in simple cases [47, 48]. The axially symmetric $g$-tensor problem is an example of where this type of calculation is practical. This approach quickly leads to extremely complicated formulas that are difficult to handle for the general case. Therefore, numerical implementation of the recursive relation Eq. 37 is essential in calculating enough coefficients of the continued fraction for an accurate approximation of $I(\Delta\omega)$. In practice, one generates the matrix representation of $\mathbf{A}$ in the $|f_j\rangle$ basis in which the resulting matrix is complex symmetric (see Section VI and Appendix A):

$$\mathbf{A}_{jk} \equiv \langle f_j | \mathbf{A} | f_k \rangle. \tag{50}$$

From Eq. 37, the standard recursive relation of the LA may be rewritten as

$$\beta_k \mathbf{q}_{k+1} = (\mathbf{A} - \alpha_k \mathbf{I}) \mathbf{q}_k - \beta_{k-1} \mathbf{q}_{k-1}, \tag{51}$$

where the column vector $\mathbf{q}_k$ consists of the components $q_{j,k}$ (cf. Eq. 42). The standard computer implementation of the complex symmetric LA [3, 4] can then be used for calculating the coefficients $\alpha_j$ and $\beta_j$ from which $I(\Delta\omega)$ can be directly calculated using the continued-fraction representation given in Eq. 49.

Finally, we must deal with the calculation of the starting vector $\mathbf{q}_1$. Given Eq. 41 for $|\Phi_1\rangle = |v\rangle$, one can obtain the components $q_{j,1}$ by computing the scalar products $\langle f_j | \Phi_1 \rangle$. This direct approach has been used frequently. However, it usually requires numerical integrations that can become unwieldy for several degrees of freedom. An alternative approach is to consider the following expression:

$$\lim_{z \to 0^+} [z\mathbf{I} + \tilde{\Gamma}] \mathbf{q}_0 = \mathbf{c}, \tag{52}$$

where $\mathbf{q}_0$ is the vector representation of $P_0^{1/2}$ and $\mathbf{c}$ is an arbitrary vector with a projection along $\mathbf{q}_0$. This follows from the fact that $P_0^{1/2}$ is the unique stationary solution of $\tilde{\Gamma}$. One can solve this equation using matrix inversion techniques or by using the conjugate gradients algorithm (cf. Section V.B). The complete starting vector $\mathbf{v}$ can be constructed by multiplying the elements of $\mathbf{q}_0$ by the appropriate spin operator terms (cf. Eq. 27).

In discussing the convergence to the correct spectrum of the continued-fraction approximant generated by the LA, it is convenient to use the following definition of the deviation $\Delta I_n$ of the true spectral lineshape $I_R(\Delta\omega)$ from that obtained after $n$ iterative steps:

$$\Delta I_n = \int_{-\infty}^{\infty} d\,\Delta\omega |I_n(\Delta\omega) - I_R(\Delta\omega)|, \tag{53}$$

where these spectral lineshape functions have been normalized to unity. It is useful to calculate the "true" converged spectral lineshape $I_R(\Delta\omega)$ using the Rutishauser diagonalization method or some other benchmark procedure whose results are not affected by the peculiar features of the LA. Using this quantity, the sufficient number of steps $n_s$ can be defined as the smallest number $n$ of Lanczos steps that assures an error $\Delta I_n$ less than the required

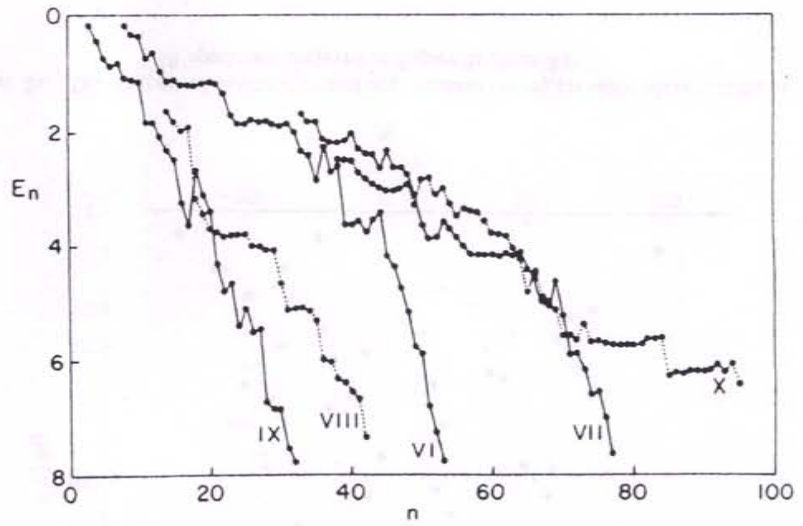
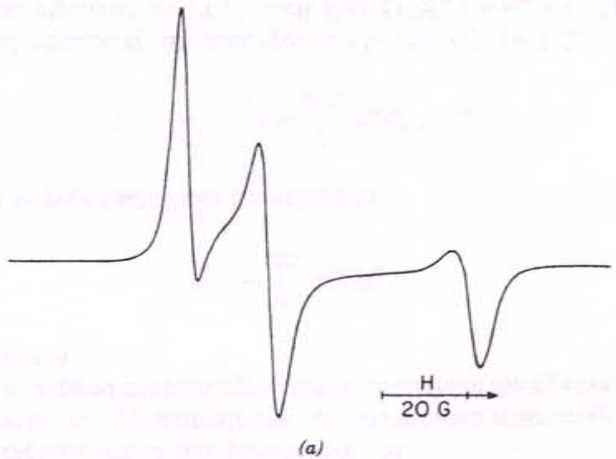



**Figure 1.** Behavior of logarithm of error: $E_n = -\log_{10}(\Delta I_n)$ as function of number of steps $n$ in LA for range of cases discussed in Ref. 3 corresponding to simple nitroxide slow-motional spectra.



accuracy for the spectral lineshape. In general, $\Delta I_n$ decreases with $n$, but it may not be strictly monotonic (cf. Figure 1 for some typical trends). For an accuracy $\Delta I_n = 10^{-4}$, $n_s$ is found to be much less than the dimension $N$ of the original matrix except in the extreme motional narrowing limit where $N$ is very small anyway. A significant portion of the computational time saved by using the LA rather than a traditional diagonalization algorithm can be attributed to this phenomenon.

The eigenvalues of $\mathbf{T}_n$ are not strictly relevant to the problem since the spectrum can be calculated directly from the elements of the tridiagonal matrix without resort to diagonalization. Therefore, any approximate form of $\mathbf{T}_n$ is adequate, independent of the effect of the loss of orthogonality of the Lanczos vectors during its generation, if it reproduces the spectral lineshape function well. However, even in this context it is still useful to analyze the eigenvalues associated with $\mathbf{T}_n$ in order to understand how the LA works for this type of problem. The slow-motional ESR spectrum displayed in Figure 2*a* has a lineshape constructed from a large collection of eigenvalues. In such cases the LA is more efficient in reproducing the overall shape of $I(\Delta\omega)$ than in exactly reproducing the eigenvalues of **A**. Figure 2*b* illustrates this fact in displaying the computed eigenvalues of the ESR problem considered in Figure 2*a*. The dots indicate the exact eigenvalues of the starting matrix **A**, which has a dimension equal to 42. The triangles represent the 16 eigenvalues of the $n_s = 16$ dimensional tridiagonal matrix whose continued-fraction approximant satisfies $\Delta I_n = 10^{-4}$. From the figure it is clear that there is no simple relation




(*a*)

**Figure 2a.** ESR absorption spectrum of hypothetical paramagnetic spin probe. Magnetic and motional parameters the same as case I in Table I of Ref. 3.

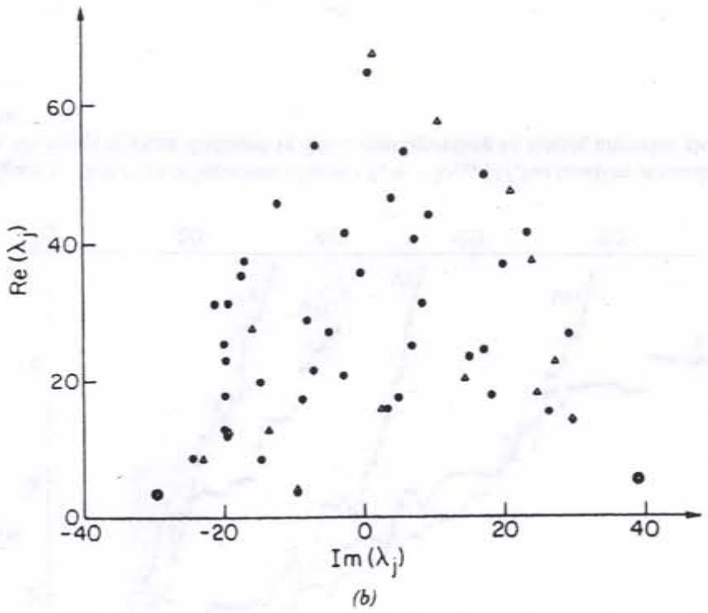



**Figure 2b.** Distribution of exact (dots) and approximate (triangles) eigenvalues relative to ESR spectrum displayed in Figure 2a (from Ref. 3).

between overall accuracy in the lineshape function and accuracy in the approximate eigenvalues. Most cannot be simply associated on a one-to-one basis with particular exact eigenvalues. Even when this is possible, the error in the approximate eigenvalues is far greater than the accuracy of $10^{-4}$ for the full spectrum. In general, the LA generates continued fractions that tend to optimize the overall shape of the spectrum rather than sets of eigenvalues. Although at first such a statement might appear contradictory, it is based on the fact that the spectral density is usually dominated by the eigenvalues of small real part, and the LA is able either to approximate them individually or to provide an "average" to a cluster of eigenvalues sufficient to represent the spectrum. In the case displayed in Figure 2*a*, one finds that only the first 11 approximate eigenvalues of lowest Re $\{\lambda_j\}$ contribute significantly, whereas of the 42 exact eigenvalues, 23 contribute significantly.

The LA approximates the spectrum surprisingly well in spite of a rather low accuracy for the eigenvalues and their components. We also see that one way to have some insight into this behavior is to look at the results in terms of an approximation of clusters of eigenvalues instead of single eigenvalues. This concept of a cluster is, of course, very qualitative. Moreover, the eigenvalues



obtained from the LA cannot be considered independent from one another in the evaluation of their effect on the spectrum. In this sense we refer to the fact that the LA produces an approximation of the optimal reduced space required to represent the spectral function instead of accurately reproducing the eigenvalues. The interpretation in terms of clusters is only a partial and qualitative explanation of this general behavior.

This tendency to approximate the optimal reduced space may be seen as follows. Let us expand the starting vector $|v\rangle$ in terms of the eigenvectors $|\Psi_m\rangle$ of the operator **A**:

$$|v\rangle = \sum_{m=1}^{N} w_m |\Psi_m\rangle \tag{54}$$

then the $l$th Krylov vector can be written as

$$|k_l\rangle = \mathbf{A}^{l-1}|v\rangle = \sum_{m=1}^{N} w_m \lambda_m^{l-1} |\Psi_m\rangle. \tag{55}$$

Thus, the $n$-dimensional subspace spanned by the $|k_l\rangle$, $l = 1, 2, \ldots, n$, cannot contain those eigenvectors $|\Psi_m\rangle$ such that $\langle v|\Psi_m\rangle = w_m = 0$. That is, the Krylov subspace, which is spanned by the Lanczos vectors, is biased to exclude the eigenvectors for which $w_m = 0$ and which therefore cannot contribute to the spectrum.

The fact that the LA only approximates the optimal reduced space is related to what we call the extreme-eigenvalue effect. Increasing the order of the subspace tends to amplify the importance of eigenvectors with larger eigenvalues $\lambda_m$ but with small $w_m$. This extreme-eigenvalue effect will be stronger the larger the magnitude of the coefficients $w_m$ for the larger eigenvalues. However, it is the eigenvectors associated with small $|\lambda_m|$ and large projections on the starting vector that are most important in the calculation of the spectrum (cf. Eq. 33). Furthermore, when finite-precision arithmetic is taken into account, a Krylov vector can have a small projection onto eigenvectors that are strictly orthogonal to the starting vector in infinite-precision arithmetic. Since these are frequently eigenvectors with large $|\lambda_m|$, the "extreme-eigenvalue effect" is likely to enhance their role in the subsequent Krylov projections. We see, in Figure 2*b*, how these effects bring in eigenvalues of large Re $\{\lambda_j\}$ even though they are found to have very small or negligible projections on the starting vector.

Therefore, the combination of the extreme-eigenvalue effect and the particular structure of the matrix generates a departure from ideal behavior of the LA (i.e., the optimal reduced space). This constitutes a negative feature of the application of the LA, but it is inherent in the method, and in spite of this,

the algorithm allows us to approximate the matrix in a subspace with dimension much smaller than the initial dimension of the matrix.

This situation can be improved by restricting the initial basis set as closely as possible, so as to exclude those basis elements that contribute appreciably to the eigenvectors with larger $|\lambda_m|$ but are unimportant in the spectrum. This is done in what follows.

Let us, however, first review the disadvantages to the use of the LA for numerical applications. Its main weakness is its loss of orthogonality among the Lanczos vectors it generates from the Krylov vectors by Gram–Schmidt orthogonalization. This is due to accumulation of numerical roundoff 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 (or Fokker–Planck spectral densities) are determined by only a small subset of eigenvectors, in particular, those associated with eigenvalues $\lambda_m$ with weakest damping (i.e., smallest Re $\{\lambda_m\}$), and approximations to these eigenvalues (or "clusters of eigenvalues") are rapidly generated such that $n_s \ll N$ for convergence to the spectrum, well before roundoff error can accumulate to the point where it can significantly affect the computation. However, roundoff 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.

Another limitation of the LA 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 roundoff 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 we referred to as the minimum truncation scheme (MTS). Knowledge of the MTS can greatly speed up calculations. In the past, such knowledge was obtained indirectly by trial-and-error calculation of spectra with different basis sets in order to specify which types of basis vectors are important. This scheme is very time consuming as well as incomplete. In previous practice, one tended to work with sets of basis vectors significantly larger than the MTS, since the latter was not convenient to determine.

We shall now consider an approach that preserves all the advantages of the LA in tridiagonalizing **A** but improves on it by (i) estimating at each recursive step the magnitude of an appropriately chosen residual so that the recursive steps can be terminated when $n = n_s$; (ii) providing a criterion to determine when computer roundoff error has become a problem; and (iii) providing a convenient means of estimating the MTS. It is based on the method of conjugate gradients discussed in the next section.



### B. The Conjugate Gradients Algorithm

The conjugate gradient (CG) method of Hestenes and Steifel [2] was originally used for solving equations of form $\mathbf{A}|u\rangle = |v\rangle$ with **A** a real symmetric positive definite (RSPD) matrix. The starting point of the CG method is to consider this equation in the form

$$|r_k\rangle = |v\rangle - \mathbf{A}|u_k\rangle, \tag{56}$$

where $|u_k\rangle$ is the $k$th approximant to $|u\rangle$, and $|r_k\rangle$ is the residual vector associated with $|u_k\rangle$. The residual vector $|r_k\rangle$ is easily seen to be the vector that gives the negative gradient of the functional $f[u_k] = \langle u_k|\mathbf{A}|u_k\rangle - \langle u_k|v\rangle$ provided **A** is RSPD, so that a minimization of $f[u_k]$ is equivalent to solving $\mathbf{A}|u\rangle = |v\rangle$. Equation 56 is solved by successive iterations that do not minimize along the sequence of vectors $|r_k\rangle$, for $k = 1, 2, \ldots, n$, which would be the method of steepest descent, but rather minimize the functional $f[u_k]$ along a set of "conjugate directions" $|p_k\rangle$ for $k = 1, 2, \ldots, n$. This procedure avoids the problem of further minimization steps spoiling the minimization along the previous conjugate direction vectors. The conjugate direction vectors $|p_k\rangle$ are defined by the equations

$$|r_{k+1}\rangle = |r_k\rangle - a_k\mathbf{A}|p_k\rangle \tag{57}$$

and

$$|p_{k+1}\rangle = |r_{k+1}\rangle + b_k|p_k\rangle. \tag{58}$$

where the $a_k$ and $b_k$ are given by

$$a_k = \frac{\langle r_k|r_k\rangle}{\langle p_k|\mathbf{A}|p_k\rangle}, \tag{59}$$

$$b_k = \frac{\langle r_{k+1}|r_{k+1}\rangle}{\langle r_k|r_k\rangle}. \tag{60}$$

The residual vectors are easily shown to be mutually orthogonal but not normalized, whereas the set of conjugate direction vectors are "A-conjugate";

that is,

$$\langle p_j|\mathbf{A}|p_k\rangle = 0 \quad \text{if} \quad j \neq k. \tag{61}$$

The vector $|p_k\rangle$ is the closest vector to $|r_k\rangle$ that is A-conjugate to the previous conjugate direction vectors. It is also true that $\langle p_j|r_k\rangle = 0$ for $j = 1, 2, \ldots, k-1$. Also, the $(k+1)$st approximant to the solution vector $|u_{k+1}\rangle$ is obtained from $|u_k\rangle$ as

$$|u_{k+1}\rangle = |u_k\rangle + a_k|p_k\rangle. \tag{62}$$

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

$$|r_1\rangle = |v\rangle - \mathbf{A}|u_1\rangle \tag{63}$$

and

$$|p_1\rangle = |r_1\rangle, \tag{64}$$

where $|u_1\rangle$ is some initial guess for $|u\rangle$. Then, for $k = 1, 2, \ldots, n-1$, one calculates, in order, $a_k$, $|u_{k+1}\rangle$, $|r_{k+1}\rangle$, $b_k$, and $|p_{k+1}\rangle$ using Eqs. 57–62. At each step the norm of the residual vector $|r_k\rangle$,

$$\|r_k\|^2 = \langle r_k|r_k\rangle, \tag{65}$$

is a measure of the deviation of $|u_k\rangle$ from the true solution $|u\rangle$.

At this stage the LA and CG appear to be very different algorithms; the former tridiagonalizes A, whereas the latter generates a sequence of approximants to the solution of Eq. 3. We consider their equivalence in the following. First we need to consider the applicability of this CG method to complex symmetric matrices. It is not hard to show that for nondefective (i.e., diagonalizable), nonsingular complex symmetric matrices the preceding CG method applies, provided only that we use the generalized norm (see Eq. 47) as in the application of the LA to complex symmetric matrices. However, for the CG method, there is the additional requirement that $\mathbf{A}'(\Delta\omega)$ be nonsingular. Recall from Section IV that the complex symmetric matrix A will have complex roots: the real parts give the linewidths and the imaginary parts, the resonance frequencies. Since each relevant eigenvalue of A must, on physical grounds, have a nonzero real part, both A and A′ will be nonsingular. Nevertheless, it is convenient to replace $i\Delta\omega$ by $i\Delta\omega + T_2'^{-1}$, where $T_2'^{-1}$ is equivalent to an additional linewidth contribution so that one avoids spurious divisions by zero that can occur when $\langle p_k|\mathbf{A}|p_k\rangle = 0$ in the calculation of $a_k$ (see Eq. 59). Zero divisors of this type can be identified as spurious through the construction of the Lanczos tridiagonal matrix from the various quantities generated by the CG algorithm, as will be shown. The invariance of the Lanczos tridiagonal matrix and the associated continued-fraction approximant to the spectral lineshape function under origin shifts are discussed in Appendix B.



However, it is useful to relax the use of the generalized norm in estimating the error in the approximate solution vector given by Eq. 56, which is used to monitor the convergence of the method. Given $|r_k\rangle$ determined from Eqs. 56 and 57, it is useful to consider two specific forms of the norm, Eq. 65, which we write in terms of its components $y_{j,k}$ of $|r_k\rangle$ in the original basis set $|f_j\rangle$ (cf. Eq. 41):

$$r_{k,\mathrm{ps}}^2 \equiv \left|\sum_j y_{k,j}^2\right|, \tag{66}$$

$$r_{k,\mathrm{H}}^2 \equiv \sum_j |y_{k,j}|^2, \tag{67}$$

whereas letting $|r_{k,\mathrm{true}}\rangle \equiv |v\rangle - \mathbf{A}'(\Delta\omega)|u_k\rangle$ (cf. Eq. 56), at each iterative step we have a third norm:

$$r_{k,\mathrm{true}}^2 = \sum_j |y_{k,j}^{\mathrm{true}}|^2. \tag{68}$$

That is $r_{k,\mathrm{H}}$ is the usual definition of a norm in unitary space, as is $r_{k,\mathrm{true}}$, whereas $r_{k,\mathrm{ps}}$ is the modulus of the generalized or pseudo norm in a complex orthogonal space.

For a complex symmetric matrix, the second and third norms are equal in exact arithmetic and are guaranteed to be real. The first norm is, however, most closely related to the generalized norm required in our CG algorithm for complex symmetric matrices. Taking the absolute value of the generalized norm of $|r_k\rangle$ gives a real value for $r_{k,\mathrm{ps}}^2$, as in Eq. 66, that can be more easily compared to the other two forms of $r^2$. In practice, it is found that $r_{\mathrm{H}}^2 = r_{\mathrm{true}}^2$ in finite-precision arithmetic until they become on the order of the unit roundoff error [80], whereas $r_{\mathrm{ps}}^2$ is always smaller. Once the value of $r_{\mathrm{H}}^2$ approaches the limit of finite-precision arithmetic, any further iterations fail to improve the quality of the approximate solution vector, and $r_{k,\mathrm{true}}^2$ remains constant as $k$ is increased. In contrast, the values of $r_{k,\mathrm{ps}}^2$ and $r_{k,\mathrm{H}}^2$ continue to decrease and therefore can no longer accurately represent the error in the approximate solution vector. However, these two forms of $r^2$ are readily available during each step, whereas $r_{k,\mathrm{true}}^2$ requires a substantial amount of extra calculation because the quantity $\mathbf{A}|u_k\rangle$ is not normally part of the CG algorithm. In light

of these facts, it is convenient to use $r^2_{k,H}$ as the criterion for convergence, and it will simply be referred to as $r^2$. On the other hand, it is advisable to occasionally compute $r^2_{k,true}$ to check if roundoff error has become a problem.

### C. The Equivalence of LA and CG Methods

To make full use of the CG algorithm, its equivalence to the LA must be recognized. First, note that the orthogonal set of vectors $|r_k\rangle$ and the conjugate set $|p_k\rangle$ are contained in the same Krylov subspace generated by **A** and $|v\rangle$, and the same is true for the Lanczos vectors [80]. Following Golub and Van Loan [80], an explicit expression for the construction of the Lanczos tridiagonal matrix $\mathbf{T}_n$ by the CG method for a RSPD matrix **A** that is valid at each stage of the iteration is

$$\mathbf{T}_k = \mathbf{D}_k^{-1}\mathbf{B}_k^{tr}\mathscr{A}\mathbf{B}_k\mathbf{D}_k^{-1}, \tag{69}$$

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

$$\mathscr{A}_{k,ii} = \langle p_i|\mathbf{A}|p_i\rangle, \qquad i = 1, 2, \ldots, k, \tag{70}$$

and **D** is also diagonal with elements

$$D_{k,ii} = \|r_i\| = \rho_i \equiv \left(\sum_j y^2_{i,j}\right)^{1/2}, \qquad i = 1, 2, \ldots, k, \tag{71}$$

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

$$B_{k,ii} = 1, \qquad B_{k,i,i+1} = -b_i, \qquad i = 1, 2, \ldots, k, \tag{72}$$

with $b_i$ given by Eq. 60. It turns out that the residual vectors are colinear with the Lanczos vectors; more precisely,

$$|\Phi_i\rangle = \pm\rho_i^{-1}|r_i\rangle, \qquad i = 1, 2, \ldots, k. \tag{73}$$

Since the direction of $|r_i\rangle$ and therefore the signs of its components are well defined by Eq. 56, the Lanczos vectors, which are normalized in an arbitrary fashion, bear the sign ambiguity since their direction is not specified by the LA. It follows from Eqs. 69–73 that the matrix elements of $\mathbf{T}_n$ are

$$\alpha_k = \langle p_k|\mathbf{A}|p_k\rangle/\rho_k^2 + (\rho_k^2/\rho_{k-1}^4)\langle p_{k-1}|\mathbf{A}|p_{k-1}\rangle, \tag{74}$$

$$\beta_k = -(\rho_k/\rho_{k-1}^3)\langle p_{k-1}|\mathbf{A}|p_{k-1}\rangle. \tag{75}$$



Thus, the elements $\alpha_k$ and $\beta_k$ of the Lanczos tridiagonal matrix $\mathbf{T}_n$ are readily obtained from quantities calculated by the CG algorithm for every step. This approach may be used to generate a tridiagonal matrix approximation to **A** using the CG algorithm in the same spirit as the LA.

Several points should be made about this equivalence. First, the tridiagonal matrices generated by the LA and CG methods are not equal. The off-diagonal matrix elements can differ in sign because of the sign ambiguity associated with the Lanczos vectors. However, the spectrum is independent of the arbitrary choice of direction of the Lanczos vectors. The change of sign of the off-diagonal matrix elements amounts to performing an equivalence transformation on the continued-fraction approximant to the spectrum (see Appendix B). In addition, one must start with $|r_1\rangle = |v\rangle$, which implies $|u_1\rangle = 0$ from Eq. 63, in order to obtain the correspondence between the Lanczos vectors and the residual vectors given by Eq. 73.

When the complications arising from the application of the CG algorithm to nondefective, nonsingular, complex symmetric matrices are considered, again it is found that this approach is applicable provided that the generalized norm and an origin shift chosen to remove spurious zero divisors are used. This means that $\rho_i$, as defined in Eq. 71, is a complex quantity.

This equivalence of the tridiagonal matrices generated by the LA and CG methods for the type of complex symmetric matrices **A** that arise in the calculation of magnetic resonance spectra has been verified by numerical calculation. It is found that the magnitudes of real and imaginary parts of the matrix elements of $\mathbf{T}_n$ obtained by the LA (Eqs. 38 and 39) and the CG algorithm (Eqs. 74 and 75) agree to at least six significant figures when double-precision arithmetic is used.

In conclusion, the CG method can be applied to complex symmetric matrices **A** to give the Lanczos tridiagonal matrix from which spectra may be calculated by the continued-fraction technique. The benefit derived from the very small amount of extra computational work necessary for the CG method as opposed to the LA method is that an objective criterion for the convergence $r^2$ can be monitored at every step. Finally, note that the basic CG algorithm 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.

### D. Minimum Truncation Scheme

As discussed previously, 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 of the continued-fraction approximants to the desired accuracy. To this end, the CG method can be used 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. 3), the knowledge of th vector $|u(\Delta\omega)\rangle$ in terms of its components $z_j(\Delta\omega) \equiv \langle f_j|u(\Delta\omega)\rangle$ in the origin basis set for a representative sample set of sweep positions should provide a accurate assessment of the importance of each basis vector in determining th spectrum.

Consider the $j$th component $z_j$. From Eq. 3 and the definition of $z_j$ it follow that

$$z_j(\Delta\omega) = \langle f_j|\mathbf{A}'^{-1}(\Delta\omega)|v\rangle = \sum_m \langle f_j|\psi_m\rangle(i\Delta\omega + \lambda_m)^{-1}\langle\psi_m|v\rangle, \quad (7$$

where, in the last equality of Eq. 76, the eigenvectors $|\psi_m\rangle$ of $\mathbf{A}$ with associate eigenvalues $\lambda_m$ were introduced. This last expression for $z_j(\Delta\omega)$ in Eq. 76 is product of three quantities. First, the scalar product $\langle\psi_m|v\rangle$, which is th projection of the $m$th eigenvector on the transition moment vector, is 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 $j$th basis vector i contributing to $|\psi_m\rangle$. These two quantities are obviously independent of th sweep variable. Finally, $(i\Delta\omega + \lambda_m)^{-1}$ expresses how the contribution of th $m$th eigenvector varies across the spectrum. If it resonates far from th frequency of the applied RF field or if it is very broad, this quantity is ver small. All these factors are needed to estimate the importance of the basi vector $|f_j\rangle$ to the spectrum. Since they are all included in $z_j(\Delta\omega)$, it can b used as a measure of the importance of $|f_j\rangle$ to the spectrum at the poin $\Delta\omega$. Provided that the spectrum is sampled at a sufficient number of value of the sweep variable, the maximum value of $|z_j(\Delta\omega)|$ over the sampled value of the sweep variable can be taken as a measure of the importance of th basis vector $|f_j\rangle$ in determining the spectrum. This may be done by solvin Eq. 3 using the CG algorithm and keeping track of the largest value o $|z_j(\Delta\omega)|$ for each basis vector as the sweep variable is varied. In practice, i is useful to monitor a slightly different function that treats the high- anc low-amplitude portions of the spectrum more equally.

This approach requires that a basis set larger than the MTS but containin the latter as a subset be utilized initially. Nevertheless, in most applications wherein calculations are compared to experimental spectra, it is necessary tc vary input parameters by small amounts 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. This makes the final search for the MTS for the larger problem less time consuming. Our examples below illustrate this.

Given that Eq. 76 is a good criterion for determining the MTS, an



alternative way to proceed would be to diagonalize the tridiagonal matrix generated by the LA by the $QR$ transform (see Section IV) but to store the Lanczos and $QR$ transformation matrices and multiply these two matrices together to obtain the projections of the eigenvectors on the various basis vectors and the starting vector. Unfortunately, it would destroy the great efficiency of the LA to keep track of the full transformation matrix. Also, for large matrices ($N \simeq 10^4$ to $N \simeq 10^5$) enormous memory may be required. Consequently, Cullum and Willoughby [89] recommended an inverse iteration procedure to obtain the important eigenvectors once a set of converged eigenvalues have been obtained by the LA. However, as stated previously, the spectra are extremely well approximated long before the actual eigenvalues have converged [3]. Thus, much more effort would be required in implementing the LA in order to achieve accurate enough eigenvalues to construct good eigenvectors by inverse iteration than is required to obtain converged spectra. On the other hand, the basic CG algorithm successfully delivers the needed information to determine the MTS.

The problem of determining the MTS has been studied using the CG method by Vasavada, Schneider, and Freed [4]. Equation 3 is simply solved using the CG algorithm for several values of the sweep variable in the range where the spectrum is expected to be nonzero. It was found that 10–20 values of $\Delta\omega$ are sufficient for slow-motional ESR spectra. In performing the sweep, it is useful to use as the initial vector (cf. Eq. 63) at the $m$th value of sweep variable, $|u(\Delta\omega)\rangle$ the solution vector from the previous value since this reduced the number of CG steps needed for convergence. This is valid because the Lanczos tridiagonal matrix need not be reconstructed to calculate the amplitude of the spectrum at a single point. In addition, the matrix $\mathbf{A}$ can be preconditioned [80] to enhance the rate of convergence [4] (see also Section V.G). The combined effects of using a good estimate for the solution vector and a preconditioning matrix that minimizes the extreme eigenvalue effect can lead to substantial savings in computational effort [4]. Also, it is sufficient for the present purpose, which is only to estimate the importance of the individual basis vectors, to use weaker convergence criteria for $r^2$. It was found that a useful measure of the overall importance of the individual basis vectors is given by

$$s_j = \max_{\Delta\omega} \left| \frac{z_j(\Delta\omega)}{\langle v|u(\Delta\omega)\rangle} \right|. \quad (77)$$

This criterion, which involves the normalization of $z_j(\Delta\omega)$ by the amplitude of the spectrum at that point, ensures that the low-amplitude regions of the spectrum are also accurately approximated. This is especially important for

two-dimensional electron-spin echo (ESE) spectroscopy, where many of the interesting variations in the contour plots occur in regions of low amplitude.

Based on these studies, it was suggested that a conservative estimate of the MTS consists of those basis vectors for which $s_j > 0.03$ for CWESR spectra and $s_j > 0.0003$ for two-dimensional ESE spectra. Of course, in preliminary calculations, one can use less stringent conditions to obtain rough approximations to the spectra.

The results of these investigations for CWESR and two-dimensional ESE are summarized in Tables 1 and 2, respectively. The initial basis sets used in these calculations were significantly larger than the MTS, as has been the normal procedure for LA calculations. A simple set of truncation rules corresponding to the maximum values of the relevant indices characterizing the basis set that are consistent with our results for the MTS were found. Using this truncation procedure yielded a basis set of dimension $N$, which is given in the tables. The actual $N_{min}$, corresponding to the dimension of the MTS for the states that satisfy the $s_j$ criterion for the particular class of calculations, is always smaller than this. Next, a search for patterns involving interrelationships between the different basis set indices that appear in the MTS was made. It was found that a slightly more complicated truncation rule involving the specification of $M_{min}$ and $M_{max}$ for each pair of $L$ and $K$ for which there were basis states satisfying the $s_j$ criterion was very effective. Using a truncation rule of this type in the form of a look-up table, the $N$-dimensional space could be reduced to $N'$, which is much closer to $N_{min}$ than is $N$. Such look-up tables are easily implemented to provide useful approximations to the MTS.

TABLE 1 Truncation Parameters and MTS for CWESR Spectra[a]

| Number | Spin Probe[b] | $R$ | $\lambda$ | $L^e_{max}$ | $L^o_{max}$ | $K_{max}$ | $M_{max}$ | $N$ | $N'$ | $N_{min}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | TEMPONE | $10^7$ | 0 | 6 | 3 | 2 | 2 | 42 | 34 | 33 |
| 2 | TEMPONE | $10^6$ | 0 | 14 | 7 | 6 | 2 | 171 | 108 | 100 |
| 3 | TEMPONE | $10^5$ | 0 | 30 | 13 | 10 | 2 | 543 | 285 | 256 |
| 4 | TEMPONE | $10^4$ | 0 | 54 | 15 | 10 | 2 | 990 | 549 | 447 |
| 5 | TEMPONE | $10^7$ | 10 | 10 | None | 2 | 2 | 63 | 26 | 26 |
| 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) | $10^7$ | 1 | 6 | 3 | 2 | 6 | 288 | 134 | 74 |
| 9 | TEMPONE (90° tilt) | $10^7$ | 10 | 10 | 9 | 4 | 4 | 822 | 129 | 69 |
| 10 | TEMPONE (90° tilt) | $10^6$ | 10 | 12 | 11 | 6 | 6 | 1779 | 533 | 245 |
| 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]Symbols: $R$ = rotational diffusion constant ($s^{-1}$); $\lambda$ = coefficient of first-order term in expansion of scaled orienting pseudopotential, $-U(\Omega)/k_BT$; $L^e_{max}$, $L^o_{max}$ = largest even value of $L$ and odd value of $L$, respectively, for which there exist basis vectors with $s_j > 0.03$, and $K_{max}$, $M_{max}$ = largest values of $K$ and $M$ for which this occurs; $N$ = dimension of matrix if *all* basis vectors whose indices are less than or equal to $L^e_{max}$, $L^o_{max}$, $K_{max}$, and $M_{max}$; $N'$ = dimension of basis set derived from lookup table based on specifying $M_{min}$ and $M_{max}$ for every important pair of $L$ and $K$ (for $\psi = 90°$ new selection rules also utilized); and $N_{min}$ = dimension of MTS (number of basis vectors for which $s_j > 0.03$).

[b]Values of $g$ and $A$ tensors for TEMPONE: $g_{xx} = 2.0088$, $g_{yy} = 2.0061$, $g_{zz} = 2.0027$, $A_{xx} = 5.8$ G, $A_{yy} = 5.8$ G, $A_{zz} = 30.8$ G. Values of $g$ and $A$ tensors for CSL: $g_{xx} = 2.0021$, $g_{yy} = 2.0089$, $g_{zz} = 2.0058$, $A_{xx} = 33.44$ G, $A_{yy} = 5.27$ G, $A_{zz} = 5.27$ G. Static magnetic field $B_0 = 3300$ G, $(\gamma_e T_2')^{-1} = 1.0$ G.



TABLE 2 Table of Truncation Parameters and MTS for Two-Dimensional ESE Spectra[a]

| Number | Spin Probe | $R$ | $\lambda$ | $L^e_{max}$ | $L^o_{max}$ | $K_{max}$ | $M_{max}$ | $N$ | $N'$ | $N_{min}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | TEMPONE | $10^7$ | 0 | 10 | 7 | 6 | 2 | 123 | 94 | 92 |
| 2 | TEMPONE | $10^6$ | 0 | 22 | 17 | 10 | 2 | 429 | 317 | 307 |
| 3 | TEMPONE | $10^5$ | 0 | 44 | 37 | 18 | 2 | 1485 | 1010 | 971 |
| 4 | TEMPONE | $10^4$ | 0 | 88 | 71 | 28 | 2 | 4614 | 2706 | 2506 |
| 5 | TEMPONE | $10^7$ | 10 | 16 | 7 | 2 | 2 | 108 | 81 | 76 |
| 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) | $10^7$ | 1 | 10 | 7 | 6 | 10 | 1440 | 752 | 586 |
| 9 | TEMPONE (90° tilt) | $10^7$ | 10 | 16 | 15 | 6 | 6 | 2601 | 931 | 607 |
| 10 | TEMPONE (90° tilt) | $10^6$ | 10 | 20 | 19 | 10 | 12 | 8196 | 3804 | 2835 |
| 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]All parameters have same meaning as in Table 1, except for $s_j$, which is taken to be 0.0003.

### E. Convergence of Lanczos–Conjugate Gradients Projections

The error $\Delta I_n$ (Eq. 53) can be used to monitor the convergence of the continued-fraction approximant to the spectral lineshape as a function of the number of Lanczos steps. This is an objective criterion, but it is impractical since it presumes one already has a converged spectrum with which to compare.

Instead, with the CG method the residual $r_0^2$ (i.e., the values of $r^2$ calculated for $\Delta\omega = 0$) is readily available at each step and may be used to determine when to terminate the CG algorithm. In general, for a complex symmetric matrix, $r_0^2$ does not decay monotonically as a function of the number of CG steps $n$, as it does for an RSPD matrix. Instead, it behaves like a damped oscillator as a function of $n$; that is, it oscillates, but its local average value decreases as $n$ increases. Typically, CWESR spectra converge very rapidly; performing CG steps until $r_0^2 \simeq 10^{-3}$ to $r_0^2 \simeq 10^{-4}$ is more than adequate. As a conservative criterion, it was suggested that the iterations be terminated when $r_0^2 \simeq 10^{-4}$, even though for most of the spectra a surprisingly low value of $r_0^2 \simeq 10^{-2}$ is sufficient. Here, $n_s$ is taken to be the value of $n$ when $r_0^2$ first reaches a value of $10^{-4}$. For several examples, the continued-fraction approximant

derived from the tridiagonal matrix produced by stopping the CG algorithm at $r_0^2 \simeq 10^{-4}$ gives $\Delta I_n \simeq 10^{-7}$ to $\Delta I_n \simeq 10^{-8}$, well within the noise range for experimental spectra.

As the motion slows down, one requires larger basis sets to adequately represent **A** (i.e., $N$ and $N_{\min}$ increase), and also there is an increase in the value of $n_s$ required to achieve the same value of $r_0^2$. This is true no matter which algorithm is utilized. However, it is observed that $n_s \ll N$ and that $n_s$ increases more slowly than $N$, so the advantage of the LA or the CG algorithm over traditional ones becomes relatively greater the slower the motions, corresponding to larger $N$. Some typical results reflecting these features are shown in Table 3. In obtaining the results in Table 3, values of $N$ significantly larger than the MTS have been used, in accordance with previous applications of the LA to the lineshape problem. However, in one case, a smaller value of $N$ that more closely corresponds to the MTS was used. In this case, $n_s$ 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 general phenomenon of how the LA seeks out an approximation to the "optimal reduced space" already discussed.

TABLE 3 Minimum Number of CG Steps and Associated Residual Values[a]

| $N$ | $r_0^2$ | $n_s$ | Case[b] | Number of Matrix Elements[c] |
|---|---|---|---|---|
| 1743 | $10^{-2}$ | 60 | A | 32917 |
| | $10^{-4}$ | 104 | | |
| | $10^{-10}$ | 172 | | |
| 429 | $10^{-2}$ | 49 | A | 7701 |
| | $10^{-4}$ | 77 | | |
| | $10^{-10}$ | 128 | | |
| 3543 | $10^{-2}$ | 76 | B | 70399 |
| | $10^{-4}$ | 159 | | |
| | $10^{-10}$ | 315 | | |
| 7503 | $10^{-2}$ | 89 | C | 288085 |
| | $10^{-4}$ | 170 | | |
| | $10^{-10}$ | 326 | | |
| 8196 | $10^{-2}$ | 57 | D | 666965 |
| | $10^{-4}$ | 80 | | |
| | $10^{-10}$ | 143 | | |

[a]Symbols: $N$ = dimension of matrix defined by $L^e_{\max}$, $L^o_{\max}$, $K_{\max}$, and $M_{\max}$ and symmetries given in Ref. 20; $r_0^2$ = residual squared (cf. Eq. 67) calculated at center of spectrum; $n_s$ = number of CG iterations.

[b]Case A: TEMPONE magnetic parameters (cf. Table 1) and $R = 10^6\,s^{-1}$. Here $N = 429$ corresponds to approximate MTS, cf. Table 2, entry 2, while $N = 1743$ represents typical larger basis set commonly utilized previously in two-dimensional ESE calculations by LA. Case B: TEMPONE magnetic parameters and $R = 10^4\,s^{-1}$. Case C: CSL magnetic parameters (cf. Table 1) and $R = 10^4\,s^{-1}$. Case D: TEMPONE magnetic parameters but with strong potential ($\lambda = 10$), $R = 10^6\,s^{-1}$, and director tilt $\psi = 90°$.

[c]Number of nonzero matrix elements in matrix counting real and imaginary parts separately.



The tendency of the LA to provide a good approximation to the spectrum before it accurately reproduces the important eigenvalues of the original matrix can be understood a little better in terms of the equivalence between the Lanczos and CG algorithms. That is, the Lanczos tridiagonalization of **A** and the solution of Eq. 3 at $\Delta\omega = 0$ using the CG algorithm involve the calculation of an orthonormal basis for the same sequence of Krylov subspaces. Thus, the LA is equivalent to the minimization of the residual to the solution vector by CG at $\Delta\omega = 0$. Furthermore, it is easy to show that the Krylov subspaces generated by $\mathbf{A}'(\Delta\omega) = \mathbf{A} + i\Delta\omega\mathbf{I}$ must be independent of $\Delta\omega$ [96]. This implies that all of the vectors generated by the CG algorithm, including the approximate solution vector for any value of $\Delta\omega$, can be expressed as linear combinations of the Lanczos vectors obtained from the LA after the same number of steps. In this manner, the equivalence of these two algorithms can be used to extend the approximate *local* solution of the problem (e.g., for $\Delta\omega = 0$) to an approximate *global* solution (i.e., for all values of $\Delta\omega$). The observation that a value of $r_0^2 \simeq 10^{-3}$ to $r_0^2 \simeq 10^{-4}$ at $\Delta\omega = 0$ is sufficient to obtain a converged spectrum for all $\Delta\omega$ is rather empirical but seems to be consistent with our observation that the outer portions (or wings) of the spectrum tend to converge sooner than the central portion for which $\Delta\omega \sim 0$. This observation is a manifestation of the extreme eigenvalue effect (cf. Eq. 55). In addition, the fact that only the projection of the approximate solution vector on the starting vector is important in determining the spectrum can be used to rationalize the relatively large error that can be tolerated in the solution vector (cf. Eq. 3). For the same reason, large errors in the approximate eigenvectors can be tolerated (cf. Eq. 33).

### F. Calculating Two-Dimensional ESE Spectra

The CG method has also been used to calculate two-dimensional ESE spectra using the approximate expression of Millhauser and Freed [30] for the two-dimensional ESE signal (cf. Section VII):

$$S(\omega,\omega') \propto \sum_j c_j^2 \frac{T_{2,j}}{1+\omega^2 T_{2,j}^2} \exp\left(-\frac{(\omega'-\omega_j)^2}{\delta^2}\right), \tag{78}$$

where for the $j$th "dynamic spin packet" (i.e., the $j$th eigenvector $|\psi_j\rangle$ of **A** corresponding to the eigenvalue $\lambda_j$), $T_{2,j}^{-1} = \mathrm{Re}\{\lambda_j\}$ is its Lorentzian width and $\omega_j = \mathrm{Im}\{\lambda_j\}$ its resonant frequency. Also, the weighting factors are given by

$$c_j^2 = \langle\psi_j|v\rangle^2 \simeq (\mathrm{Re}\langle\psi_j|v\rangle)^2, \tag{79}$$

where the approximate equality is valid only in the very slow-motional region where Eq. 78 is approximately valid. The two-dimensional ESE spectrum is

inhomogeneously broadened (with respect to the $\omega'$ sweep variable) by convolution with a Gaussian distribution of half-width $\delta$. For purposes of testing the computational method, Eq. 78 was utilized with the approximate form for $c_j^2$ in Eq. 79 even when the motion was too fast for these expressions to accurately represent an experimentally obtainable spectrum.

It is clear from Eq. 78 that it is necessary to obtain estimates of the eigenvalues $\lambda_j$ that contribute with nonnegligible weight factors. This was done by diagonalizing the tridiagonal matrix $\mathbf{T}_n$ derived from the CG method utilizing standard procedures [3, 45]. The approximate eigenvectors $|\psi_j\rangle$ 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 [12, 45] (see also Section IV).

Because the two-dimensional ESE spectra require significantly more accurate estimates of the eigenvalues and the weights, the 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 CWESR spectrum. In particular, we find that $r_0^2 \simeq 10^{-8}$ to $r_0^2 \simeq 10^{-10}$ is sufficient for two-dimentional ESE spectra, 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 3.

In order to study the convergence further, it is possible to introduce, by analogy to Eq. 53, the following definition of the error in the two-dimensional ESE spectrum:

$$\Delta S_n = \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} d\omega' |S_n(\omega, \omega') - S_R(\omega, \omega')|, \tag{80}$$

where $S_R(\omega, \omega')$ is the normalized "exact two-dimensional ESE spectrum" from a complete diagonalization of $\mathbf{A}$ or some very good approximation to it, and $S_n(\omega, \omega')$ is the normalized approximate spectrum obtained by diagonalizing $\mathbf{T}_n$. The examples studied indicate that $r_0^2 \simeq 10^{-10}$ corresponds to $\Delta S_n \simeq 0.007$ to $\Delta S_n \simeq 0.02$.

Studies of the MTS for two-dimensional ESE spectra (cf. Table 2) showed that it is necessary to retain basis vectors for which $s_j \simeq 0.0006$ to $s_j \simeq 0.0003$, so a cutoff of $s_j = 0.0003$ was recommended for determining the MTS for two-dimensional ESE rather than the $s_j = 0.03$ value recommended for CWESR. It is interesting that a more stringent application of the same criterion (Eq. 77) that was appropriate for CWESR problems is also useful for two-dimensional ESE calculations, which require substantially greater accuracy due to their greater sensitivity to the approximate eigenvalues. This matter is further discussed by Vasavada, Schneider, and Freed [4].



### G. Direct Calculation of Spectra and Spectral Densities by Conjugate Gradients

In the preceding section it has been shown how the CG algorithm can be employed to calculate the CWESR spectrum for several values of the sweep variable in order to determine the MTS and the minimum number of CG or Lanczos steps. It is, of course, possible to employ the CG algorithm to directly compute the entire CWESR spectrum from Eq. 3. Preliminary studies show that the calculation of the magnitude of the spectrum at one value of the sweep variable is, on the average, about five times faster than doing enough CG steps to get the entire spectrum from the tridiagonal matrix. This implies that a direct calculation using the CG algorithm for an entire CWESR spectrum using 200 values of the sweep variable would take about 40 times longer than the LA.

One might hope that if the CG calculation is performed for smaller increments of the sweep variable, only a few iterations per sweep position would be required if the solution at the previous sweep position is used as an initial guess for the following point. Numerical experiments show that if the number of sweep positions is increased by a factor of 10 from 20 (for the MTS calculation) to 200, the computer time required to complete the calculation increases by only a factor of 5.5 for termination at $r^2 = 10^{-4}$ and by only a factor of 3.6 for $r^2 = 10^{-2}$. This indicates that one does improve the efficiency of the calculation in this manner, especially if a relatively large $r^2$ is sufficient. Further improvements in efficiency might also be achieved if the increment in the sweep variable was chosen in an adaptive manner.

Another way to speed up the direct calculation is by preconditioning the $\mathbf{A}'(\Delta\omega)$ matrix, as mentioned previously. Preconditioning is a general device to improve the convergence of an iterative solution to a matrix problem (e.g., Eq. 3). Given an RSPD 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 also RSPD [80]. In the case of complex symmetric matrices, if one can let $\mathbf{M} = \mathbf{\Gamma} + T_2'^{-1}\mathbf{I}$ and $\mathbf{N} = i(\mathbf{L} - \Delta\omega\mathbf{I})$, $\mathbf{M}$ can be made symmetric and is positive definite. The preconditioned problem one solves is

$$(\mathbf{M}^{-1/2}\mathbf{A}'\mathbf{M}^{-1/2})\mathbf{M}^{1/2}|u\rangle = \mathbf{M}^{-1/2}|v\rangle = \mathbf{A}''|u'\rangle = |v'\rangle. \tag{81}$$

For isotropic Brownian rotational diffusion, where the eigenvalues of the diffusion operator are proportional to $L(L+1)$, the stochastic Liouville matrix $\mathbf{A}'$ 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

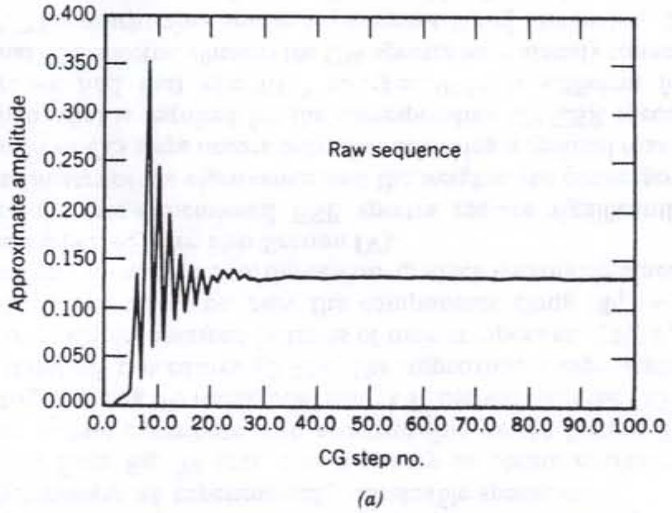
diagonal elements and to scale the other matrix elements accordingly. For more general cases, such as when a restoring potential is needed, it is useful to keep the calculation of **M** as simple as possible. Taking **M** to be a diagonal matrix comprised of the diagonal elements of $\mathbf{\Gamma} + T_2'^{-1}\mathbf{I}$ in these instances seems to work fine. In general, we find that the preconditioned CG algorithm does speed up the convergence of the calculation. Unfortunately, because preconditioning is not a similarity transformation, it cannot be used in a simple way to diagonalize **A** for purposes of calculating spectra. For an example of where this is possible see Chapter 6.

Based on these results, it is clear that tridiagonalization by the LA is the more efficient method for CWESR, and the direct method does not even apply to two-dimensional ESE using the formula given in Eq. 78. However, there are some cases where the LA method is not as suitable as the direct CG method. These include:

- CWESR of transition metal ions, which require a wide range of sweep of the static magnetic field [17].
- CWESR in the presence of strong saturating radiation fields [97, 99].
- Calculating the effect of finite amplitude–finite length RF pulses on a spin system.

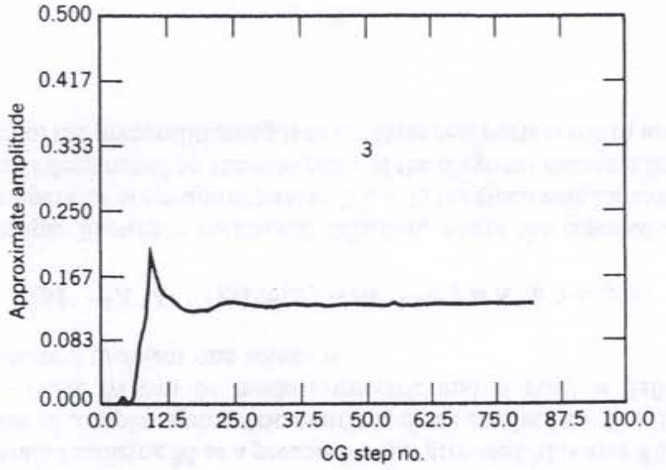
In all cases, the elements of the matrix $\mathbf{A}'(\Delta\omega)$ have nontrivial dependence on the sweep variable.

Further enhancement of the speed and efficiency of the direct CG method can be made by acceleration of the convergence by extrapolation methods. We

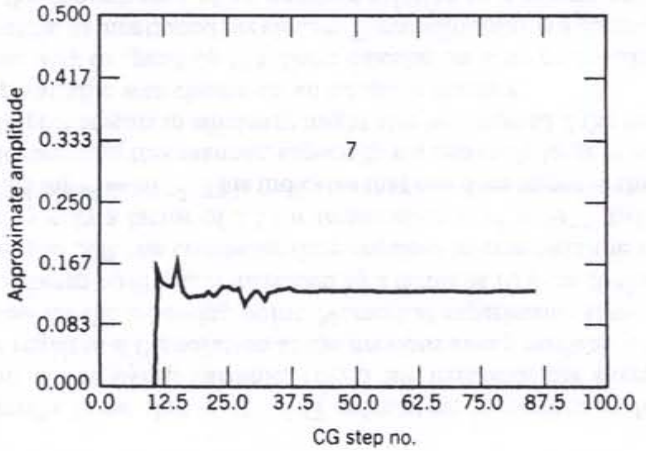



*(a)*






*(b)*




*(c)*

**Figure 3.** Effects of number of terms included in the $\varepsilon$ extrapolation of the approximate values of the amplitude of the spectrum. This calculation is for slow-motional nitroxide spectrum at $\Delta\omega = 0$: (*a*) raw data; (*b*) result of three-term $\varepsilon$ extrapolation; (*c*) result of seven-term $\varepsilon$ extrapolation.

may expect that later recursive steps merely remove undesirable "transients" in the sequence of approximate values of the spectrum. An effective means of improving the convergence rate of a sequence of this type is the Shanks transformation, or Padé extrapolation, or epsilon expansion. A preliminary study has shown that the Padé extrapolation method can be effective in

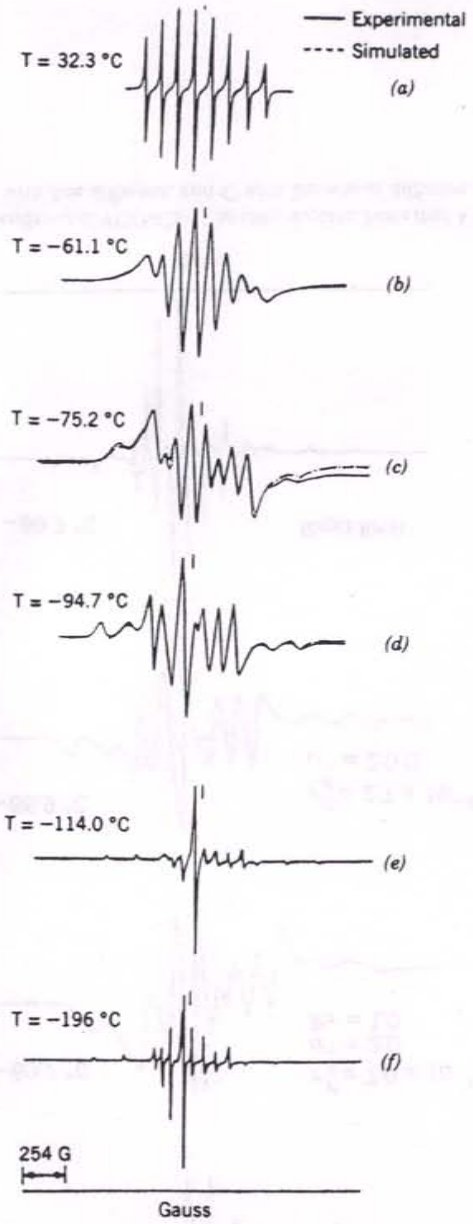



**Figure 4.** Comparison of experimental and simulated spectra from the fast motional to th rigid limit for VO(acac$_2$(pn)) in toluene. All calculations use a Brownian rotational diffusio model: (*a*) $\tau_R = 2.06 \times 10^{-11}$ s; (*b*) $\tau_R = 2.63 \times 10^{-11}$ s; (*c*) $\tau_R = 5.00 \times 10^{-10}$ s; (*d*) $\tau_R = 2.25 \times 10^{-10}$ s; (*e*) $\tau_R = 5.00 \times 10^{-8}$ s; (*f*) rigid limit. (From Ref. 17. Reprinted with permissio from R. F. Campbell and J. H. Freed, *J. Phys. Chem.* **84**, 2668. Copyright 1980 America Chemical Society).



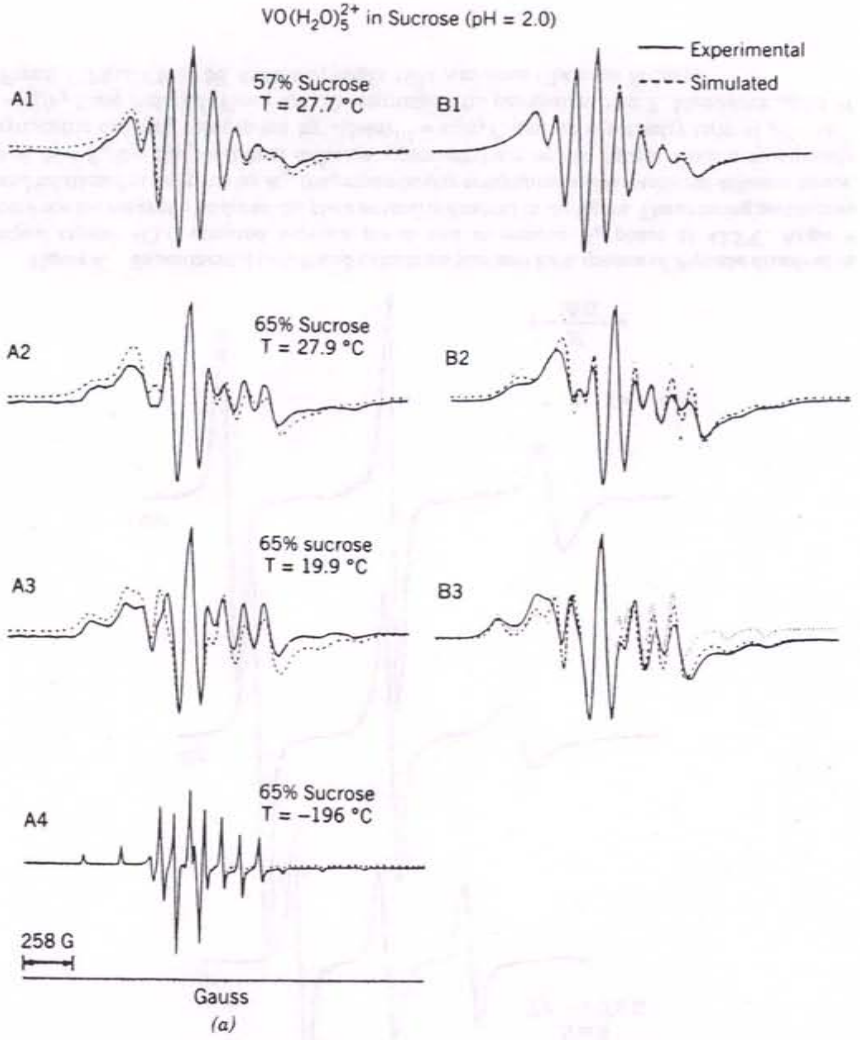



**Figure 5a.** Model dependence of $VO(H_2O)_5^{2+}$ in sucrose, Series A: Comparison of experiment with moderate-jump diffusion. Series B: Comparison of moderate-jump diffusion (solid lines) with its free (dashed lines) and Brownian (dotted lines) diffusion equivalent. (Moderate jump gave best agreement in all cases.) (A1) $\tau_R^{jump} = 3.4 \times 10^{-10}$ s; (A2) $\tau_R^{jump} = 6.0 \times 10^{-10}$ s; (A3) $\tau_R^{jump} = 9.0 \times 10^{-10}$ s; (A4) rigid limit. (From Ref. 17. Reprinted with permission from R. F. Campbell and J. H. Freed, *J. Phys. Chem.* **84**, 2668. Copyright 1980 American Chemical Society).

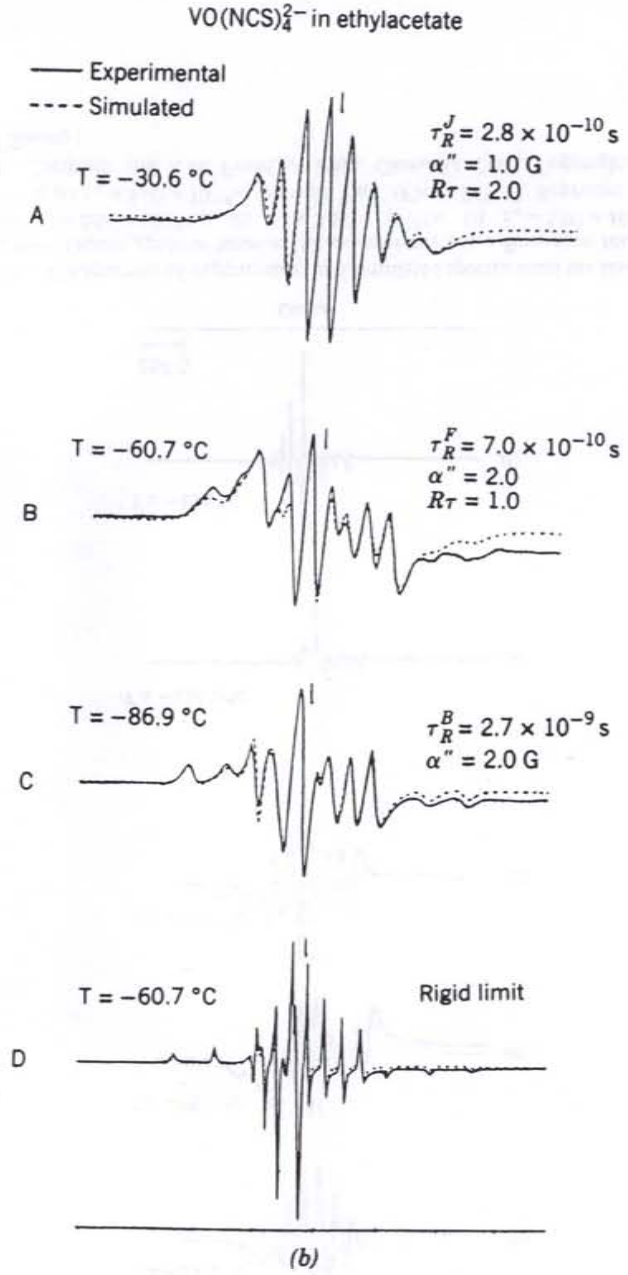



**Figure 5b.** Model dependence of VO(NCS)$_4^{2-}$ in ethyl acetate. Note that A is approximately fit with moderate jump, B with free diffusion, and C with Brownian diffusion. (From Ref. 17.)





reducing the amplitude of the rapid variations in the estimated magnitude of the spectrum that occur for the first few CG steps (cf. Figure 3). It is possible that this could be used to reduce the number of CG steps required to give an accurate estimate of the value of the spectrum for a given value of the sweep variable [4]. More details on these techniques and their relationship to the continued-fraction approximants for the spectrum are given in Appendix B. Although the full power of such methods has not been fully explored, they do show substantial promise.

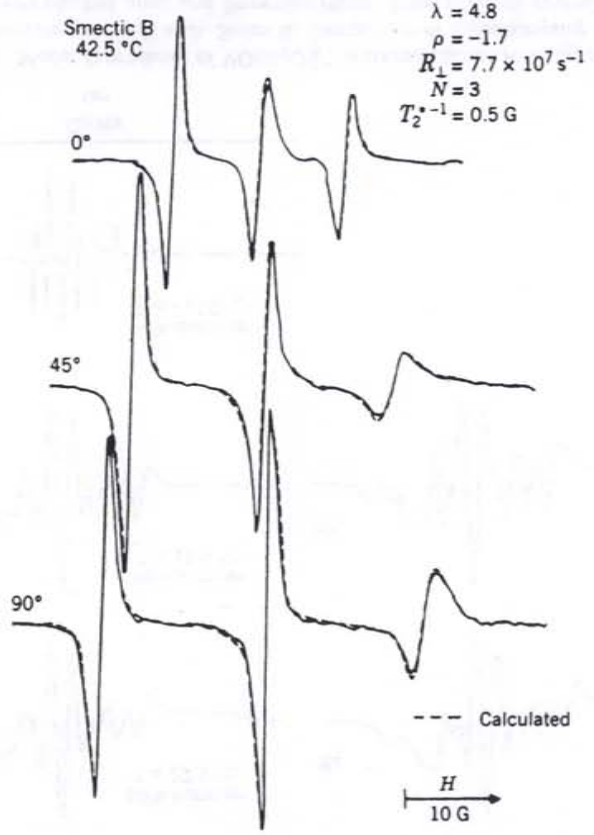



**Figure 6.** Experimental (solid) and calculated (dashed) ESR spectra of P-probe dissolved in liquid crystal 4O,6 oriented between plates and in smectic $B_A$ phase at 42.5°C. Angle $\theta$ between the magnetic field and the plate normal is denoted in the figure. The ordering parameters and rotational rates (given by $R_\perp$, the perpendicular component of the rotational diffusion tensor, and $N = R_\parallel/R_\perp$, the rotational diffusion asymmetry) are on the figures. Both a cylindrically symmetric ordering term, given by $\lambda(5/4\pi)^{1/2} = \varepsilon_0^2/k_B T$, and an asymmetry term of $\rho(5/4\pi)^{1/2} = \varepsilon_2^2/k_B T$, are included (From Ref. 20. Reprinted with permission from E. Meirovitch and J. H. Freed, *J. Phys. Chem.* **88**, 4995. Copyright 1984 American Chemical Society).

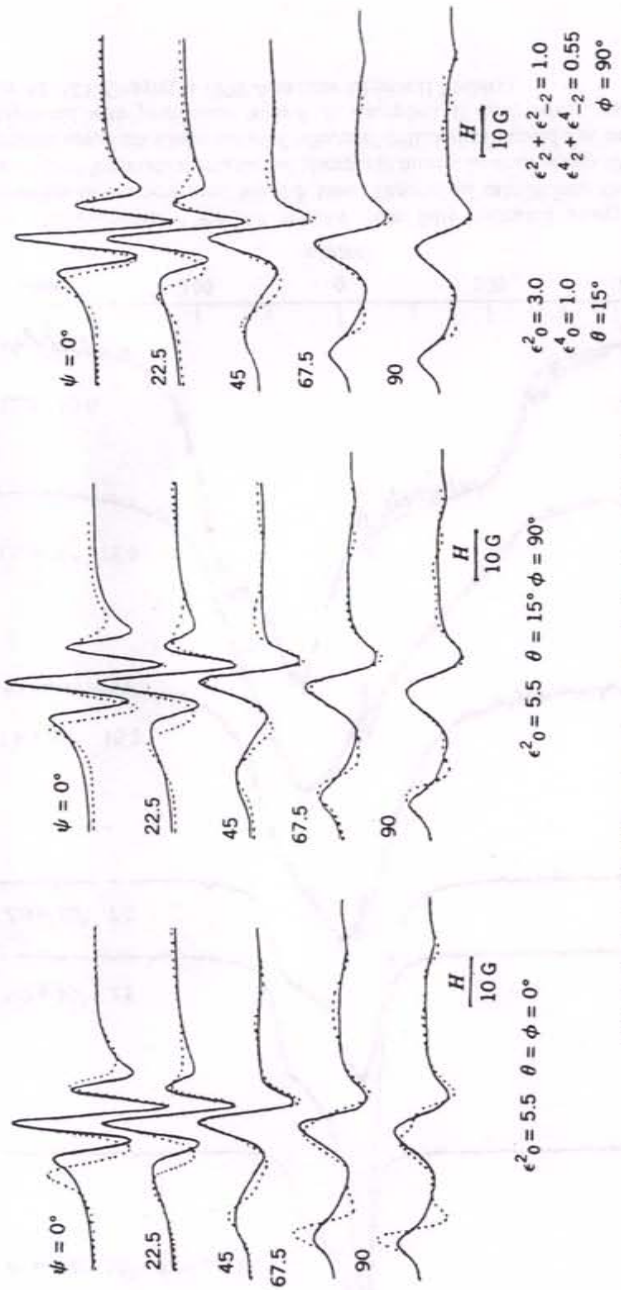



**Figure 7.** Experimental spectra of homeotropically aligned CSL in liquid crystal S2 at −8 °C for tilt angle $\psi$ between the liquid crystal director and the static magnetic field (solid lines). Dashed lines: Simulated spectra with anisotropic diffusion in a high ordering potential. $(0, \theta, \phi)$ denote the Euler angles between the magnetic frame and the ordering potential frame. The $\epsilon^L_K$ denote coefficients in the expansion of the scaled ordering potential in spherical harmonics. (From Ref. 21. Reprinted with permission from E. Meirovitch and J. H. Freed, *J. Phys. Chem.* **88**, 4995. Copyright 1984 American Chemical Society).





## H. Slow-Motional ESR Spectra: Examples

We have already pointed out that slow-motional spectra provide considerably more information about the microscopic models of rotational dynamics than motionally narrowed spectra. Thus, for example, jump models of rotational reorientation lead to slow-motional spectra that are distinguishable from Brownian reorientational models [12–14]. Whereas nitroxide-type spin probes with $^{14}N$ nuclear spin of $I = 1$ are commonly studied [12], previous work has shown that $VO_2^+$ complexes ($^{51}V$ nuclear spin of $I = 7/2$) are more sensitive to the choice of motional model. Also, the latter exhibit slow motions for $\tau_R \simeq 100$ ps, as compared to 1 ns for nitroxides at normal 9-GHz ($X$-band) frequencies. In particular, the slow-tumbling lineshapes seem to be strongly dependent on the nature of the ligands and of the solvent such that a range of different models had to be used (cf. Figures 4 and 5*a*, *b*).

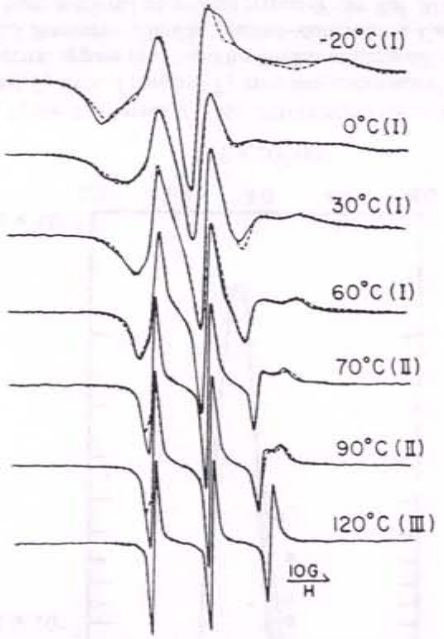



**Figure 8.** ESR spectra from macroscopically unoriented lipid (DPPC dispersions) with 7 wt % water containing nitroxide end labeled lipid. Solid lines: Experimental spectra. Dashed lines: Calculations based on model of microscopically ordered lipid fragments macroscopically disordered (MOMD model). Phase I: Low-temperature gel phase. Phase II: Liquid crystalline $L_\alpha$ phase. Phase III: New high-temperature phase exhibiting residual microscopic ordering. Typical order parameters: 0.33, 0.18, and 0.02 in Phases I, II, and III, respectively. Typical values of $R_\perp$, perpendicular rotational diffusion coefficient: of $1.0 \times 10^{-8}$, $10 \times 10^{-8}$, and $20 \times 10^{-8}\,s^{-1}$, respectively, in these phases. (From Ref. 23. Reprinted with permission from H. Tanaka and J. H. Freed, *J. Phys. Chem.* **88**, 6633. Copyright 1984 American Chemical Society).

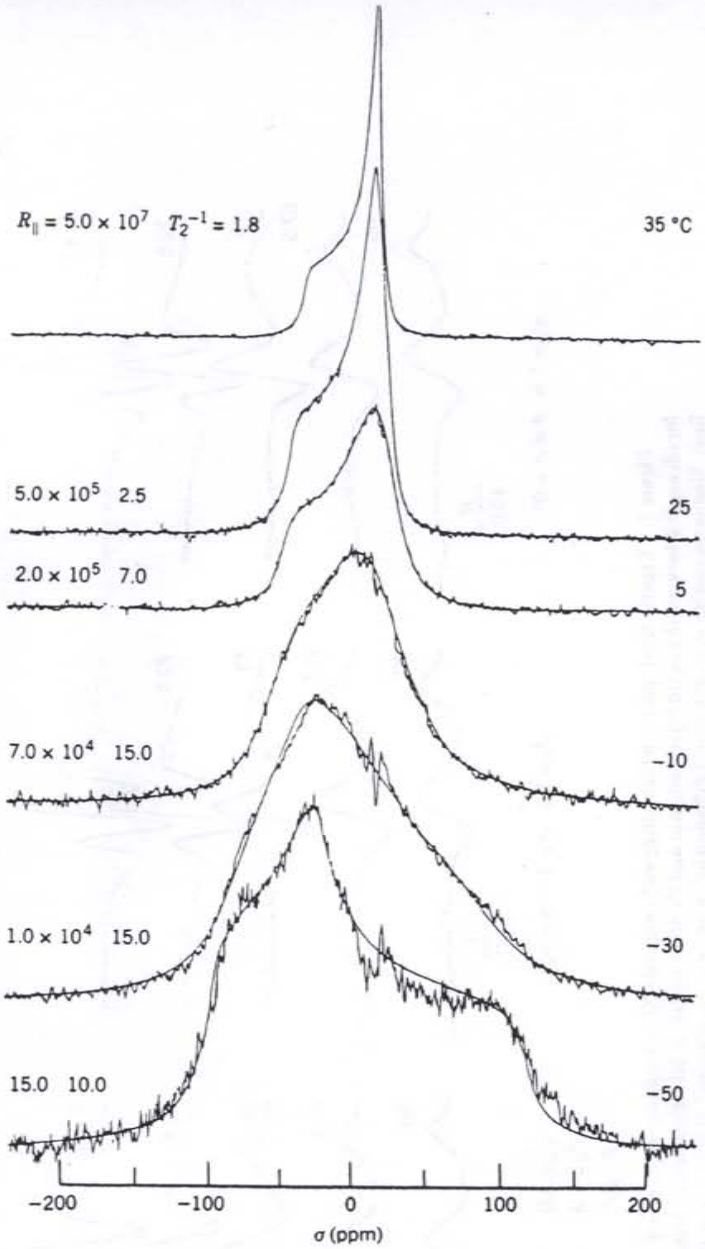



**Figure 9.** Experimental $^{31}$P NMR spectra from fully hydrated phospholipid (DPPC) bilayers for range of temperatures. Smooth lines: Theoretical calculations showing rotational diffusion coefficient for internal rotation of phosphate moiety in head group. Overall rotation is too slow to affect spectrum. Orientation of "effective" diffusion axis could also be specified. (From Ref. 36. Reprinted with permission from R. F. Campbell, E. Meirovitch, and J. H. Freed, *J. Phys. Chem.* **83**, 525. Copyright 1979 American Chemical Society).



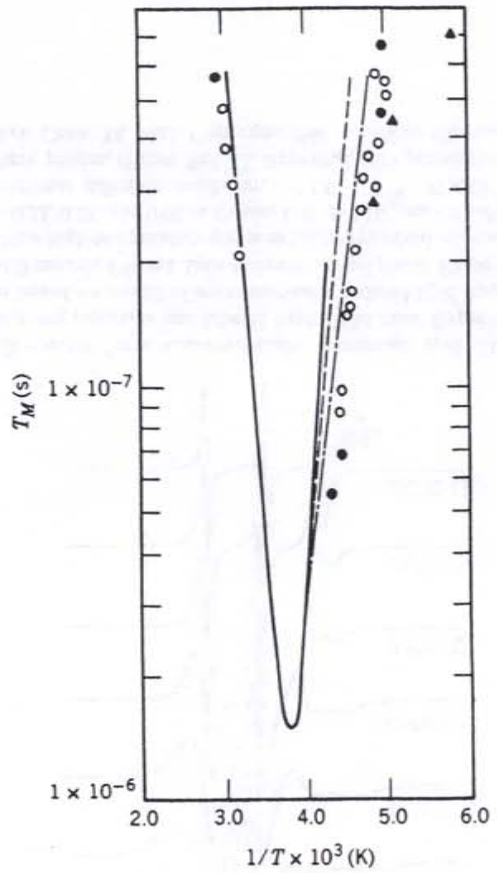



**Figure 10.** Graph of $T_2$ (or $T_M$) versus $T^{-1}$ for PD-TEMPONE in 85% glycerol to 15% water. Circles: conventional $T_2$ data. Triangles: $T_2$ from two-dimensional ESE. Different lines represent $T_2$ in central spectral regime calculated for models of jump diffusion (solid line), free diffusion (dashed line), and Brownian diffusion (dashed–dotted line). Calculations employed values of $\tau_R$ extrapolated from motional narrowing regime (From Ref. 30.)

In liquid crystalline phases the most important new feature of molecular rotational dynamics is that it must occur relative to a mean orienting potential. In Figure 6 we show spectra of a nitroxide-labeled liquid crystallike probe molecule in a low-temperature smectic-B phase. The spectrum was fit by a nonspherically symmetric diffusion tensor and a noncylindrically symmetric orienting potential. Orienting potentials in uniaxial liquid crystals may generally be expanded in spherical harmonics, but most experiments provide insufficient information to evaluate beyond the $L = 2$ spherical harmonics. In

TABLE 4 Summary of Rotational Relaxation Mechanisms

| | Mechanism | Characteristic | Parameters | References |
|---|---|---|---|---|
| 1. | Anisotropic diffusion | Unequal reorientation rates about principal axis system fixed in molecule; when reorientation rates are equal, this becomes isotropic Brownian diffusion. | $R_\parallel = (6\tau_{R_\parallel})^{-1}$ rotational diffusion coefficients about molecular symmetry axis; $R_\perp = (6\tau_{R_\perp})^{-1}$ rotational diffusion coefficient about molecular axes perpendicular to symmetry axis; $N = \tau_{R_\perp}/\tau_{R_\parallel}$ | 5, 12, 13, 16, 149 |
| 2. | Anisotropic viscosity | Unequal reorientation rates about principal axis system fixed in laboratory (e.g., dc magnetic field axis) | $\hat{R}_\parallel = (6\hat{\tau}_{R_\parallel})^{-1}$ rotational diffusion coefficient about orienting axis in lab frame; $\hat{R}_\perp = (6\hat{\tau}_R)^{-1}$ rotational diffusion coefficient perpendicular to orienting axis in lab frame; $\tilde{N} = \hat{\tau}_{R_\perp}/\hat{\tau}_{R_\parallel}$ | 9, 14, 16, 149, 155, 161 |
| 3. | Fluctuating torques | Anisotropic torques that induce reorientation are themselves relaxing at rate that is not much faster than reorientation of probe molecule | $\tau_R$ = relaxation time for rotational diffusion of probe molecule; $\tau_M$ = relaxation time for fluctuating torques inducing reorientation of probe molecule; $\varepsilon \equiv (1 + \tau_M/\tau_R)^2$; may be combined with analogues of models 1 (requiring additional specification of $\tau_{M_\parallel}$ and $\tau_{M_\perp}$) or 2 (requiring $\hat{\tau}_{M_\parallel}$ and $\hat{\tau}_{M_\perp}$) | 14, 16, 41, 153, 161 |
| 4. | SRLS | Probe molecule relaxes in local potential field, while latter averages out isotropically at significantly slower rate than rate of probe molecule | $S_l = 1/2\,(3\cos^2\theta - 1)$, order parameter of probe relative to local anisotropic potential field (i.e., local structure); $\tau_x$ = relaxation rate of local structure (may also be combined with analogues | 14, 16, 41, 142, 149, 156, 161, 215 |
| 5. | Jump diffusion | Molecule reorients by random jumps of arbitrary angle | $R = (\varepsilon^2)_{av}/6\tau$, where $\tau$ = time between jumps and $(\varepsilon^2)_{av}$ = mean-square jump angle; can be generalized to include anisotropic diffusion tensors by analogy with models 1 or 2 | 12–14, 17 |
| 6. | Free diffusion | Molecular reorientation is partially due to free gaslike motion that is perturbed by frictional effects of surroundings | $\tau_R$ and $\tau_J$, where $\tau_J$ = angular momentum relaxation time; they are often related by the Hubbard–Einstein relation: $\tau_R\tau_J = \mathcal{I}/6kT$, $\mathcal{I}$ = moment of intertia; also, $\tau_J^{-1} = \beta$, friction coefficient | 12, 15 |
| 7. | Director fluctuations | Hydrodynamic effect in which nematic director fluctuates in its orientation with respect to applied magnetic field | $\tau_q^{-1} = Kq^2/\eta$, where $\tau_q$ = relaxation time of $q$th Fourier component of director fluctuation, with $K$ = average elastic constant of liquid crystal and $\eta$ = viscosity | 16, 128 |
| 8. | Critical fluctuations | Pretransitional effects near (almost) second-order phase transition; leads to apparent divergences in spin relaxation due to divergences in correlation length and/or relaxation time for order parameter | $\tau_q^{-1} = \omega_\xi(1 + q^2\xi^2)^x$, where, for isotropic–nematic (nematic–smectic) transitions, $\omega_\xi = L/v\xi$, with $L$ = force constant for order fluctuations and $v$ = associated viscosity ($\omega_\xi = \tau_m^{-1}$, with $\tau_m$ = relaxation time of smectic clusters), $\xi$ = coherence length of nematic (smectic) order fluctuations and $x = 1(\frac{3}{4})$ | 128, 141, 150–152 |
| 9. | Discrete Jumps | Molecule reorients by discrete jumps among equivalent sites; often this reorientation is about single internal axis | $\tau_{dj}$ = mean time between jumps | 40, 213, 214 |

Reprinted with permission from J. S. Hwang, K. V. S. Rao, and J. H. Freed, *J. Phys. Chem.* **80**, 1970. Copyright 1976 American Chemical Society. (This is an up-dated version).

Figure 7, we illustrate, with an oriented spectrum from a large and rigid nitroxide-labeled molecule, one of the exceptional cases where there is sufficient sensitivity to detect influence of the $L = 4$ spherical harmonics as well.

In spin label studies on oriented lipid multilayers it was possible to discern three transitions, the main [gel to $L_\alpha(1)$ liquid crystalline] transition, a new transition from $L_\alpha(1)$ to a very weakly ordered phase, and finally the transition to an isotropic phase at the highest temperatures. All these phases could be characterized in terms of molecular ordering and dynamics (cf. Figure 8). Thus, the main transition was characterized as primarily a "chain diffusional" transition, whereas the new transition is primarily a "chain orientational" transition. That is, there is a larger relative change in the ordering of the spin probe at the second transition, whereas the diffusion coefficient for the chain labels experiences a more significant relative change at the main transition. These are just a few examples of the kind of information that one can obtain from the slow-motional ESR studies. A related NMR study is shown in Figure 9. As a result of a number of such studies [113], a variety of rotational motional models have been inferred [114]. They are summarized in Table 4.

A clear demonstration of the implications of slow motion may be obtained from ESR spin–echo experiments that directly measure $T_2$. An echo experiment performed on a nitroxide spin probe in viscous solvent shows that (cf. Figure 10), at high temperatures corresponding to the motional narrowing region, $T_2 \propto \tau_R^{-1}$ (the rotational correlation time), whereas at low temperatures corresponding to the slow-motional regime, $T_2 \propto \tau_R^{\alpha}$, where $1/2 \leqslant \alpha \leqslant 1$. The value of $\alpha$ depends on the model for reorientation (e.g., $\alpha = 1$ for jump diffusion and $\alpha = 1/2$ for Brownian motion). This leads to a $T_2$ minimum at the appropriate midrange temperature.

## VI. THE FOKKER–PLANCK APPROACH TO MODELING MOLECULAR DYNAMICS

In the Fokker–Planck approach to modeling the molecular dynamics of a complicated many-body system, the exact Newtonian or quantum-mechanical equations of motion for all of the particles in the system are replaced by a much simpler equation only involving the relevant dynamical variables (i.e., those that are relevant for the description of some set of experimental data) and the coupling of these variables to a stochastic thermal bath representing the effects of the neglected variables. The rationale for this procedure is rather well known from a physical point of view for systems at or very near to thermal equilibrium, but the rigorous mathematical justification is quite complicated, especially for systems far from equilibrium [108, 115–121]. The discussion here is mainly directed toward the use of Fokker–Planck operators to model the angular position, angular velocity, and translational degrees of freedom of spin probes in condensed phases and on surfaces for the purpose of interpreting ESR spectra, though the range of applicability of the material presented here extends far beyond this relatively narrow topic.

The reduction of the Fokker–Planck operator for a planar rotator, including both angular position and velocity, to complex symmetric matrix form is carried out in detail as a simple example of the practical use of the theoretical and mathematical concepts presented here.

The correlation functions of interest are of the form

$$g(t) = \overline{f_1^*(t) f_2(0)} \tag{82}$$

$$= \langle f_1 | e^{-\Gamma t} P_0 | f_2 \rangle, \tag{83}$$

where the overbar indicates averaging over the canonical ensemble, and $f_1$ and $f_2$ are arbitrary functions of the dynamical variables and external forces. The associated spectral functions are the Fourier–Laplace transforms of the correlation functions

$$\hat{g}(z) = \int_0^\infty e^{-zt} g(t)\, dt \tag{84}$$

$$= \int_0^\infty \langle f_1 | e^{-(z\mathbf{I} + \Gamma)t} P_0 | f_2 \rangle\, dt. \tag{85}$$

The observable spectral densities are given by taking the limit of Eq. 85 as the real part of the complex frequency variable $z = \delta + i\omega$ approaches zero from positive values:

$$\hat{g}(i\omega) = \lim_{\delta \to 0^+} \langle f_1 | [(\delta + i\omega)\mathbf{I} + \Gamma]^{-1} P_0 | f_2 \rangle. \tag{86}$$

This equation is formally equivalent to Eq. 1.

### A. General Fokker–Planck Operators

The general multivariate time-independent nonlinear Fokker–Planck operator can be written as [115, 119]

$$\Gamma(q, \lambda) = \sum_{i=1}^{N} \frac{\partial}{\partial q_i} \left\{ K_i(q, \lambda) - \frac{1}{2} \sum_{j=1}^{N} \frac{\partial}{\partial q_j} K_{ij}(q, \lambda) \right\}, \tag{87}$$

where $q$ is an $N$-dimensional real vector of classical stochastic dynamical

variables and $\lambda$ represents the parametric dependence on a set of classical external fields. The term nonlinear here refers to the fact that the real functions $K_i(q,\lambda)$ are not necessarily linear in the dynamical variables and the real functions $K_{ij}(q,\lambda)$ are not constants independent of $q$. The differential equations for the various probability distributions (Eqs. 88, 89, and 91) are all linear partial differential equations. The set of variables contained in $q$ are assumed to constitute a stationary Markov process. It will also be useful to assume that each of the dynamical variables and external fields have well-defined parity with respect to the operation of (classical) motion reversal (see Section VI.B). The drift coefficients $K_i(q,\lambda)$ and the diffusion coefficients $K_{ij}(q,\lambda) = K_{ji}(q,\lambda)$ do not explicitly depend on time but may depend on the external fields. In modeling the motions of spin probes in liquids and liquid crystals, the external field dependence of the diffusion coefficients can be neglected.

The Fokker–Planck operator determines the time evolution of the conditional probability $P(q'|q;\lambda,t)$ via the forward Kolmogorov equation

$$\frac{dP(q'|q;\lambda,t)}{dt} = -\Gamma(q',\lambda)P(q'|q;\lambda,t) \tag{88}$$

and its adjoint, the backward Kolmogrov equation,

$$\frac{dP(q'|q;\lambda,t)}{dt} = -\Gamma^{\dagger}(q,\lambda)P(q'|q;\lambda,t), \tag{89}$$

both subject to the same initial condition

$$P(q'|q;\lambda,0) = \delta(q'-q). \tag{90}$$

It has been assumed here that natural boundary conditions have been enforced to define a unique solution (i.e., enforcing periodicity in angular variables and/or the vanishing of the solution and its derivative for large values of the coordinates and velocities). The conditional probability must satisfy Eqs. 88 and 89 for all positive times.

The Fokker–Planck equation for the time evolution of the probability density,

$$\frac{\partial P(q,t;\lambda)}{\partial t} = -\Gamma(q,\lambda)P(q,t;\lambda), \tag{91}$$

can be derived from Eq. 88 by multiplying both sides by $P(q',t';\lambda)$ followed by integration over $q'$ and the use of the stationarity properties.



The stationary probability distribution $P_0(q,\lambda)$ is determined by the unique solution to

$$\Gamma(q,\lambda)P_0(q,\lambda) = 0. \tag{92}$$

In most cases discussed in this review, $P_0(q,\lambda)$ is just the equilibrium probability distribution.

### B. Time Reversal, Detailed Balance, and Symmetrized Fokker–Planck Operators

The invariance of the fundamental microscopic equations of motion under time reversal is one of the central features of both classical and quantum mechanics. The time reversal invariance of the microscopic equations is closely related to the conditions of detailed balance for transitions between macroscopic states characterized by macroscopic observables. Making the connection between the microscopic dynamics and the macroscopic observables is one of the central problems in statistical mechanics.

A symmetrized form of a general Fokker–Planck operator that satisfies certain detailed balance restrictions is derived in this section [122]. The behavior of the eigenfunctions of the symmetrized Fokker–Planck operator under the combined operation of complex conjugation and classical motion reversal suggest that a set of basis functions that are invariant under this operation will simplify practical calculations. The general scheme for the use of basis sets of this type in forming a finite-dimensional complex symmetric matrix approximation to the symmetrized Fokker–Planck operator for computational work is discussed in Section VI.B.1, and an example of the application of the general method to the planar rotator problem is carried out in Section VI.B.3. The use of the LA for calculating spectral functions is dealt with in Section VI.B.2. The presentation given here regarding detailed balance and its implications relies heavily on the work by van Kampen [115], Graham and Haken [123], Risken [122, 119], Haken [116] and Lax [124, 125].

In the Hamiltonian formulation of classical mechanics the state of a system is described as a point in phase space, and the time evolution is given by Hamilton's equations for the coordinates $x_k$ and momenta $p_k$,

$$\frac{dx_k}{dt} = \frac{\partial H(x,p)}{\partial p_k}, \tag{93}$$

$$\frac{dp_k}{dt} = -\frac{\partial H(x,p)}{\partial q_k}, \tag{94}$$

for $k = 1, 2, \ldots, N$, where $H(x,p)$ is the Hamiltonian function for the system, which is assumed to be time independent, and is a quadratic function of the

momenta. Under these assumptions, the form of Eqs. 93 and 94 is invariant under the substitutions

$$t \to -t, \qquad x \to x, \qquad p \to -p. \tag{95}$$

This is the classical version of time reversal invariance of the microscopic equations of motion for a closed, isolated physical system. This invariance implies that if the system is at the point $(x', p')$ in phase space at time $t'$ and the equations of motion carry this point into a point $(x'', p'')$ at some later time $t''$, the same equations of motion predict that if the system had started out at the point $(x'', -p'')$ at time $t'$, it would be carried into the point $(x', -p')$ at time $t''$.

If external forces such as a magnetic field are present, the previous discussion must be modified, since the Hamiltonian is no longer a quadratic function of the momenta. Instead, a more general relation involving the reversal of both the momenta and external fields is needed. In this context it is important to clearly define the operation of motion reversal. The definition of the motion reversal operation used here is the reversal of the momenta of all the particles in the system as well as the currents giving rise to the external fields while leaving the coordinates of the particles in the system unchanged. The macroscopic dynamical variables and fields were previously assumed to have well-defined parity with respect to motion reversal. Let $\tilde{q}$ and $\tilde{\lambda}$ be the sets of motion-reversed variables and fields, respectively. The components of these vectors are given by $\tilde{q}_i = \varepsilon_i q_i$ and $\tilde{\lambda}_j = \nu_j \lambda_j$, where $\varepsilon, \nu = \pm 1$. It is also convenient to introduce an *antilinear* classical time reversal operator $\mathcal{T}_c$, which is the combined operation of complex conjugation and motion reversal:

$$\mathcal{T}_c[\gamma f(q, \lambda)] = \gamma^* f^*(\tilde{q}, \tilde{\lambda}), \tag{96}$$

where $\gamma$ is a complex constant and $f(q, \lambda)$ is an arbitrary function of $\lambda$. As pointed out by Lax [124], it is unnecessary to include the complex conjugation operation under some circumstances, but it is required here.

The transition to the macroscopic scale can be accomplished by defining a set of macroscopic dynamical variables $q$ that are functions of the microscopic variables $x$ and $p$ followed by a suitable averaging over the equilibrium distribution in phase space. A careful analysis [115] shows that the time reversal symmetry of the microscopic equations leads to the detailed balance relation for the transition rates given by Eqs. 97 and 99.

The definition of detailed balance for a stationary process in the presence of external fields is that the conditional and stationary probabilities satisfy [115, 116]

$$P(q'|q; t, \lambda) P_0(q, \lambda) = P(\tilde{q}|\tilde{q}'; t, \tilde{\lambda}) P_0(\tilde{q}', \tilde{\lambda}). \tag{97}$$



An alternative formulation of this definition can be derived by evaluating the derivative with respect to time of both sides of Eq. 97 at $t = 0$ to get a relationship satisfied by the transition probability defined as

$$w(q', q; \lambda) = \left. \frac{dP(q'|q; t, \lambda)}{dt} \right|_{t=0} \tag{98}$$

and the stationary distribution. It is

$$w(q', q; \lambda) P_0(q, \lambda) = w(\tilde{q}, \tilde{q}'; \tilde{\lambda}) P_0(\tilde{q}', \tilde{\lambda}). \tag{99}$$

It should be noted that these conditions (Eqs. 97 and 99) must be satisfied for all pairs of dynamical variables $q'$ and $q$ independently. More complete derivations of the relationship between macroscopic detailed balance and microscopic time reversal symmetry are given by van Kampen [115] and Lax [124].

It is physically reasonable and can easily be shown that the stationary distribution is invariant under motion reversal,

$$P_0(q, \lambda) = P_0(\tilde{q}, \tilde{\lambda}). \tag{100}$$

The stationary probability distribution can be used to define a real-valued function $\Phi(q, \lambda)$ that can be interpreted as a generalized thermodynamic potential function,

$$P_0(q, \lambda) = N e^{-\Phi(q, \lambda)}, \tag{101}$$

where $N$ is a normalization factor and the potential function must satisfy $\Phi(q, \lambda) = \Phi(\tilde{q}, \tilde{\lambda})$ since $P_0$ has this property.

It is now possible to derive a set of conditions that must be satisfied by the drift and diffusion coefficients and the stationary distribution such that the detailed balance conditions given in Eqs. 97 and 99 hold. Since the functions $K_i(q, \lambda)$ are arbitrary well-behaved functions of $q$ and $\lambda$, the drift terms in Eq. 87 do not necessarily have simple transformation properties under motion reversal. However, it is straightforward to define linear combinations of the drift coefficients and their motion-reversed counterparts,

$$D_i(q, \lambda) \equiv \tfrac{1}{2}[K_i(q, \lambda) + \varepsilon_i K_i(\tilde{q}, \tilde{\lambda})], \tag{102}$$

$$J_i(q, \lambda) \equiv \tfrac{1}{2}[K_i(q, \lambda) - \varepsilon_i K_i(\tilde{q}, \tilde{\lambda})], \tag{103}$$

to simplify the analysis. The drift terms in Eq. 87 can now be expressed in a

form in which the reversible or irreversible nature of the individual terms is made clear.

The insertion of the form of the general Fokker–Planck operator (Eq. 87) into the Fokker–Planck equation (Eq. 91) shows that the Fokker–Planck equation has the form of the divergence of a probability current $S$,

$$\frac{\partial P(q,t;\lambda)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial S_i(q,t;\lambda)}{\partial q_i}, \tag{104}$$

where the components of $S$ are given by

$$S_i(q,t;\lambda) = \left\{ K_i(q,\lambda) - \frac{1}{2}\frac{\partial}{\partial q_j} K_{ij}(q,\lambda) \right\} P(q,t;\lambda). \tag{105}$$

The components of the probability currents can be broken down into irreversible and reversible parts using the associated definitions for the drift coefficients (Eqs. 102 and 103). The irreversible and reversible components of the probability current are, respectively,

$$S_i^{(+)}(q,t;\lambda) \equiv D_i(q,\lambda) - \frac{\partial}{\partial q_j}[K_{ij}(q,\lambda)P(q,t;\lambda)], \tag{106}$$

$$S_i^{(-)}(q,t;\lambda) \equiv J_i(q,\lambda)P(q,t;\lambda). \tag{107}$$

Equation 92 defining the stationary distribution can now be rewritten as the vanishing of the divergence of the probability current $S_0(q,\lambda)$ associated with $P_0(q,\lambda)$. Inserting $P_0(q,\lambda)$ into Eqs. 106 and 107 give

$$S_{0i}^{(+)}(q,\lambda) = D_i(q,\lambda)P_0(q,\lambda) - \frac{\partial}{\partial q_j}[K_{i,j}(q,\lambda)P_0(q,\lambda)]. \tag{108}$$

$$S_{0i}^{(-)}(q,\lambda) = J_i(q,\lambda)P_0(q,\lambda), \tag{109}$$

Requiring that the divergence of $S_0$ vanish and using the known symmetry of $P_0(q,\lambda)$ with respect to motion reversal gives the following necessary and sufficient conditions for detailed balance to hold [116, 119]:

$$K_{ij}(q,\lambda) = \varepsilon_i\varepsilon_j K_{ij}(\tilde{q},\tilde{\lambda}), \tag{110}$$

$$\sum_j \frac{\partial}{\partial q_i}[J_i(q,\lambda)P_0(q,\lambda)] = 0, \tag{111}$$

$$D_i(q,\lambda)P_0(q,\lambda) - \sum_j \frac{\partial}{\partial q_j}[K_{i,j}(q,\lambda)P_0(q,\lambda)] = 0. \tag{112}$$



The vanishing of the divergence of the stationary probability current does not mean that the probability current itself must vanish, even in the stationary or equilibrium state.

The definition of the transition probability (Eq. 98) and the forward and backward Kolmogorov equations (Eqs. 88 and 89) at $t = 0$ are very closely related. In fact, inserting the appropriate initial conditions (Eq. 90) and using the definition of the transition probability, it follows that [119, 122]

$$\Gamma(q',\lambda)\delta(q'-q)P_0(q,\lambda) = \Gamma^{\dagger}(\tilde{q}',\tilde{\lambda})\delta(\tilde{q}'-\tilde{q})P_0(q',\lambda). \tag{113}$$

Note that there are two sets of free variables in this equation, $q'$ and $q$. Using the properties of the delta function, it follows that $\delta(q'-q) = \delta(\tilde{q}'-\tilde{q})$, and it is possible to replace $q$ for $q'$ in the argument of $P(q',\lambda)$ on the left side to give

$$[\Gamma(q',\lambda)P_0(q',\lambda) - P_0(q',\lambda)\Gamma^{\dagger}(\tilde{q}',\tilde{\lambda})]\delta(q'-q) = 0. \tag{114}$$

Since Eq. 114 must hold for any arbitrary value of $q$, it is equivalent to the *operator* equation [116, 122]

$$\Gamma(q',\lambda)P_0(q',\lambda) = P_0(q',\lambda)\Gamma^{\dagger}(\tilde{q}',\tilde{\lambda}). \tag{115}$$

In this equation, $P_0(q',\lambda)$ is treated as an operator in the same way a potential function is treated as an operator in quantum mechanics. It must be stressed that Eq. 115 is an *operator* equation that must be valid for operating on arbitrary functions. The derivation of Eq. 115 is based on the original presentation of Risken [122] and the review by Haken [116]. For alternative derivations of this important result based on a master equation, see the book by Risken [119] and the paper by Lax [124].

The operator equation 115 can be put into a more symmetric form by pre- and postmultiplication by the operator $P_0^{-1/2}(q',\lambda)$ to give

$$P_0^{-1/2}(q',\lambda)\Gamma(q',\lambda)P_0^{1/2}(q',\lambda) = P_0^{1/2}(q',\lambda)\Gamma^{\dagger}(\tilde{q}',\tilde{\lambda})P_0^{-1/2}(q',\lambda). \tag{116}$$

The so-called symmetrized Fokker–Planck operator $\tilde{\Gamma}(q,\lambda)$ defined as

$$\tilde{\Gamma}(q',\lambda) = P_0^{-1/2}(q',\lambda)\Gamma(q',\lambda)P_0^{1/2}(q',\lambda) \tag{117}$$

then satisfies

$$\tilde{\Gamma}(q',\lambda) = \tilde{\Gamma}^{\dagger}(\tilde{q}',\tilde{\lambda}). \tag{118}$$

This symmetrized form of the Fokker–Planck operator will play a central role

in the derivation of matrix element selection rules and the calculation of spectral functions using the complex symmetric LA.

The special symmetry of $\tilde{\Gamma}(q,\lambda)$ is also reflected in its eigenvectors. When $\tilde{\Gamma}$ contains both reversible and irreversible terms, the Hermitian and anti-Hermitian parts of $\tilde{\Gamma}$ do not, in general, commute and the eigenvectors of $\tilde{\Gamma}(q,\lambda)$ and $\tilde{\Gamma}^{\dagger}(q,\lambda)$ are not the same. However, Eq. 118 can be used in conjunction with the usual biorthogonality relations [126–129] to derive a physically important correspondence between members of the two sets of eigenvectors [116, 122].

Let the $\{\Psi_i(q,\lambda)\}$ be the set of eigenfunctions of $\tilde{\Gamma}(q,\lambda)$ with eigenvalues $\{a_i\}$ and the set $\{\Psi^j(q,\lambda)\}$ be the eigenfunctions of $\tilde{\Gamma}^{\dagger}(q,\lambda)$ with associated eigenvalues $\{b^j\}$,

$$\tilde{\Gamma}(q,\lambda)\Psi_i(q,\lambda) = a_i\Psi_i(q,\lambda), \tag{119}$$

$$\tilde{\Gamma}^{\dagger}(q,\lambda)\Psi^j(q,\lambda) = b^j\Psi^j(q,\lambda). \tag{120}$$

The biorthogonality relation is easy to derive by examining matrix elements of $\tilde{\Gamma}$ between eigenfunctions $\Psi^j$ and $\Psi_i$ [129]. Using the definition of the adjoint operator and Eqs. 119 and 120, it is clear that

$$\int \Psi^{j*}(q,\lambda)[\tilde{\Gamma}(q,\lambda)\Psi_i(q,\lambda)]\,dq = a_i\int \Psi^{j*}(q,\lambda)\Psi_i(q,\lambda)\,dq \tag{121}$$

$$= \int [\tilde{\Gamma}^{\dagger}(q,\lambda)\Psi^j(q,\lambda)]^*\Psi_i(q,\lambda)\,dq \tag{122}$$

$$= b^{j*}\int \Psi^{j*}(q,\lambda)\Psi_i(q,\lambda)\,dq. \tag{123}$$

Taking the difference between Eqs. 121 and 123 gives

$$(a_i - b^{j*})\int \Psi^{j*}(q,\lambda)\Psi_i(q,\lambda)\,dq = 0, \tag{124}$$

which implies that

$$\int \Psi^{j*}(q,\lambda)\Psi_i(q,\lambda)\,dq = 0 \tag{125}$$

if $a_i \neq b^{j*}$. Carrying out the same steps with the adjoints of Eqs. 119 and 120 shows that the sets of eigenvalues must be complex conjugates of one another; that is $a_i = b^{j*}$ for some pair of indices $i, j$. A relabeling of the eigenfunctions



and eigenvalues can then be carried out such that the usual form of the biorthogonality relation

$$\int \Psi^{j*}(q,\lambda)\Psi_i(q,\lambda)\,dq = \delta_{ij} \tag{126}$$

is valid. If any of the eigenvalues are degenerate and the sets of eigenfunctions are complete, an argument similar to the one used in quantum mechanics can be used here to show that the eigenfunctions corresponding to degenerate eigenvalues can also be chosen such that they are orthogonal.

The derivation of the general biorthogonality relation (Eq. 126) does not take advantage of the known symmetry of $\tilde{\Gamma}(q,\lambda)$ under motion reversal (Eq. 118). The implications of this symmetry are made most evident by looking at the motion-reversed counterpart of Eq. 120,

$$\begin{aligned}\tilde{\Gamma}^{\dagger}(\tilde{q},\tilde{\lambda})\Psi^j(\tilde{q},\tilde{\lambda}) &= \tilde{\Gamma}(q,\lambda)\Psi^j(\tilde{q},\tilde{\lambda}) \\ &= b^j\Psi^j(\tilde{q},\tilde{\lambda}) \\ &= a_j^*\Psi^j(\tilde{q},\tilde{\lambda}).\end{aligned} \tag{127}$$

Equations 119 and 127, and the fact that $\tilde{\Gamma}(q,\lambda)$ is invariant under complex conjugation imply that

$$\Psi_i^*(\tilde{q},\tilde{\lambda}) = \Psi^i(q,\lambda); \tag{128}$$

that is, biorthogonal partners are simply related by the classical time reversal operator. Inserting this result into Eq. 126 shows that an eigenfunction of $\tilde{\Gamma}(q,\lambda)$ must be orthogonal to the motion reversed counterparts of all the other eigenfunctions,

$$\int \Psi_j(\tilde{q},\tilde{\lambda})\Psi_i(q,\lambda)\,dq = \delta_{ij}. \tag{129}$$

If the eigenfunctions $\{\Psi_i(q,\lambda)\}$ are expanded in a complete orthonormal set of basis functions $\{\phi_i(q,\lambda)\}$ that are invariant under the classical time reversal operation,

$$\phi_i(q,\lambda) = \mathcal{T}_c\phi_i(q,\lambda) = \phi_i^*(\tilde{q},\tilde{\lambda}), \tag{130}$$

then Eq. 129 can be reexpressed in terms of the components of the eigenfunctions in this basis,

$$\sum_k c_i^k c_j^k = \delta_{ij}, \tag{131}$$

where $c_i^k = \int \phi_k^*(q,\lambda)\Psi_i(q,\lambda)\,dq$. Thus, the eigenvectors are not orthogonal in the usual unitary or Hilbert space sense. Instead, vectors orthogonal in the sense of Eq. 131 are called rectanormal rather than orthonormal [90] (see Appendices A and C). Basis functions obeying Eq. 130 can be constructed from an arbitrary complete orthonormal set by projecting the linear combinations that are even or odd with respect to motion reversal and multiplying the latter by $i$.

In writing Eq. 131 we have assumed for simplicity that any external fields exhibit motion reversal (i.e. $\tilde{\lambda}_j = \lambda_j$), and/or $\tilde{\Gamma}(q,\lambda)$ (cf. Eq. 118) is invariant with respect to $\lambda \to \tilde{\lambda}$. More generally, as Lax [125] has shown, a generalized motion reversal operator may be introduced (e.g. by multiplying by a two-fold rotation in space) so that the field (e.g. a magnetic field) is invariant to this generalized motion reversal. We assume below that the condition(s) necessary for Eq. 131 to be valid are fulfilled.

A matrix **C** constructed from the components of the eigenvectors ($C_{ij} = c_i^j$) in a basis of this type must be a complex orthogonal matrix, $\mathbf{C}^{tr}\mathbf{C} = \mathbf{I}$, by Eq. 131. This matrix, by its construction, is the transformation matrix that diagonalizes the matrix **A** of the operator $\tilde{\Gamma}$,

$$\mathbf{C}^{tr}\mathbf{A}\mathbf{C} = \mathrm{diag}(a_1, a_2, \ldots). \tag{132}$$

Since **A** is diagonalized by a complex orthogonal transformation, it must be a complex symmetric matrix in the basis $\{\phi_i(q,\lambda)\}$. It is extremely important to realize the fact that **A** is complex and symmetric. This is a consequence of the following:

- Symmetry of $\tilde{\Gamma}(q,\lambda)$ (Eq. 118).
- Choice of basis set adapted to symmetry of $\tilde{\Gamma}$ (Eq. 130).
- Invariance of operators $\tilde{\Gamma}(q,\lambda)$ and $\tilde{\Gamma}^{\dagger}(q,\lambda)$ under complex conjugation (Eqs. 87, 101, and 117).

There may also be bases that do not satisfy Eq. 130 that also render **A** complex symmetric. This is usually traceable to additional symmetries under which $\tilde{\Gamma}(q,\lambda)$ is invariant. An example of this phenomenon is given in Section VI.B.3. Thus, the validity of Eq. 130 is a sufficient but not necessary condition for the matrix of $\tilde{\Gamma}$ to be complex symmetric. The properties of complex symmetric matrices, complex orthogonal matrices, and rectanormal vectors are studied in greater detail in Appendices A and C.

In concluding, we may mention that, in general, any matrix representation of $\Gamma$ may be transformed to complex symmetric form **A** (representing $\tilde{\Gamma}$) by a similarity transformation **S**, which can be written in "polar form" by $\mathbf{S} = \mathbf{U}\mathbf{H}$, where **U** is a unitary operator and **H** is a positive definite Hermitian operator [87]. As a result of the above symmetries, it is immediately possible to identify



**H** as $P_0^{-1/2}$ and **U** as the transformation from an arbitrary orthonormal basis set to one obeying Eq. 130. Then the special properties of complex *orthogonal* spaces will apply to **A** and **C**. Note, however, that the matrix forming **A** and the scalar products forming **C** are computed in the context of the appropriate *unitary* (or Hilbert) space in which the operators $\Gamma$ and $\tilde{\Gamma}$ are defined (e.g., Eq. 149 or Eq. 279 below). This duality will be important in the analysis of time correlation functions (cf. Eq. 133) given in Section 6.B.2.

### 1. *Reduction of General Fokker–Planck Operators to Complex Symmetric Matrix Form*

The reduction of a general multivariate Fokker–Planck operator to a finite dimensional complex symmetric matrix form is a prerequisite for the implementation of the complex symmetric LA for the calculation of classical correlation functions [3, 43]. A finite dimensional matrix approximation to the symmetrized Fokker–Planck operator is easily obtained by the method of expansion in a finite set of basis functions that are invariant under the classical time reversal operation as outlined in Section VI.B. This discretization method is very convenient because it allows one to choose basis functions for the expansion in which the matrix of the symmetrized Fokker–Planck operator is automatically complex symmetric.

It is useful to make a few comments on the nature of the approximations made in this discretization process. The first comment is related to the necessity of identifying a physically relevant time scale for the process of interest, and the second is related to the close relationship between the method of moments in Hilbert space and the LA.

First, it is important to keep in mind the spirit of approximations initially used in constructing the Fokker–Planck operator. In particular, only the relevant dynamical variables with *long* relaxation times are included in the set $q$ (see also Section VI.C). The remaining variables, whose relaxation times are much shorter, are not explicitly included in the Fokker–Planck description of the stochastic process. Thus, a physically reasonable time scale must be identified to even write down an appropriate Fokker–Planck operator. The mathematical nature of the general Fokker–Planck operator is such that it may have an infinite number of eigenvalues and eigenfunctions. Furthermore, the real parts of the eigenvalues, which correspond to the relaxation rates of the corresponding eigenfunctions, may be arbitrarily large. The inclusion of these rapidly decaying modes is somewhat unsatisfactory since the dynamics of these modes may be affected by the time dependence of some of the dynamical variables that were neglected in the construction of the Fokker–Planck operator. Thus, a finite dimensional approximation to the Fokker–Planck operator may be obtained by only using the projection of the Fokker–Planck operator on the subspace spanned by the slowly relaxing eigenfunctions. The approximation derived in this manner is consistent with the spirit of the

physical approximations needed to use a Fokker–Planck operator to model the dynamics.

The difficulty in implementing this projection scheme is that if the eigenvalues and eigenfunctions of the full Fokker–Planck operator are known, the problem has already been solved! The resolution of this tautology is to use physical arguments to identify a set of basis functions whose span should contain the subspace of slow modes. This process is often simplified by considering a closely related problem whose solution is known analytically, the use of asymptotic approximations, or considering the Krylov vectors generated by $\tilde{\Gamma}$ and $|P_0^{1/2} f_2\rangle$ or $\tilde{\Gamma}^\dagger$ and $|P_0^{1/2} f_1\rangle$ (cf. Eq. 137a below). Once a set of basis functions has been chosen, it is easy to generate a finite dimensional approximation to $\tilde{\Gamma}$ by simply evaluating matrix elements.

### 2. *Calculation of Classical Time Correlation Functions and Spectral Densities with the Complex Symmetric Lanczos Algorithm*

In order to utilize the complex symmetric LA to calculate spectral functions, it is necessary to generate a finite dimensional complex symmetric matrix approximation to the symmetrized Fokker–Planck operator $\tilde{\Gamma}$ (see Eq. 117). This reduction is discussed in Sections VI.B and VI.B.1. The general forms of the correlation function (Eq. 83) and spectral function (Eq. 85) can be reexpressed in terms of the symmetrized Fokker–Planck operator (Eq. 117),

$$g(t) = \langle P_0^{1/2} f_1 | e^{-\tilde{\Gamma} t} | P_0^{1/2} f_2 \rangle, \tag{133}$$

$$\hat{g}(z) = \langle P_0^{1/2} f_1 | [z\mathbf{I} + \tilde{\Gamma}]^{-1} | P_0^{1/2} f_2 \rangle. \tag{134}$$

The complex symmetric LA is applicable in cases where the components of $|P_0^{1/2} f_1\rangle$ are the complex conjugates of the components of $|P_0^{1/2} f_2\rangle$ in the basis chosen in which the matrix of $\tilde{\Gamma}$ is complex symmetric. This is really not a restriction for the general application of the complex symmetric LA, as will be shown here.

Let $\{\phi_n(q, \lambda), n = 1, 2, \ldots, N\}$ be a set of orthonormal basis functions that are invariant under the classical time reversal operation and span the subspace of slowly relaxing eigenfunctions of $\tilde{\Gamma}$. The general symmetrized spectral function can be expressed in terms of the expansion in this basis as

$$\hat{g}(z) = \sum_{i,j=1}^{N} \langle P_0^{1/2} f_1 | \phi_i \rangle \langle \phi_i | [z\mathbf{I} + \tilde{\Gamma}]^{-1} | \phi_j \rangle \langle \phi_j | P_0^{1/2} f_2 \rangle. \tag{135}$$

When rewritten in matrix–vector notation, this reads

$$\hat{g}(z) = \mathbf{u}^\dagger [z\mathbf{I} + \mathbf{A}_N]^{-1} \mathbf{v}, \tag{136}$$



where $u_i = \langle \phi_i | P_0^{1/2} f_1 \rangle$, $v_j = \langle \phi_j | P_0^{1/2} f_2 \rangle$, and $(A_N)_{ij} = \langle \phi_i | \tilde{\Gamma} | \phi_j \rangle$.

Since the functions $f_1(q, \lambda)$ and $f_2(q, \lambda)$ are arbitrary functions, the components of $\mathbf{u}$ and $\mathbf{v}$ are, in general, complex. The simplest case to study is where $f_1(q, \lambda)$ and $f_2(q, \lambda)$ are simply related by the classical time reversal operator. If

$$f_2(q, \lambda) = \alpha \mathcal{T}_c f_1(q, \lambda) \tag{137}$$

for some complex constant $\alpha$, the vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent. In this case, the sequence of Krylov vectors generated by successive application of $\mathbf{A}_N$ on $\mathbf{v}$ span the same subspace as the Krylov vectors generated by $\mathbf{A}_N^\dagger$ and $\mathbf{u}$. Because of this fact, the complex symmetric LA can be used directly to calculate spectral functions that satisfy Eq. 137. It is important to note that spectral functions corresponding to autocorrelation functions where the components of $f_1$ are real in the basis chosen fall into this category.

On the other hand, if $f_1(q, \lambda)$ and $f_2(q, \lambda)$ are not related by Eq. 137, $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, the Krylov subspaces generated by $\mathbf{A}_N$ and $\mathbf{v}$ and by $\mathbf{A}_N^\dagger$ and $\mathbf{u}$ are inequivalent, and the basic complex symmetric LA cannot be used without further work. There is a straightforward way to resolve this difficulty. The complex symmetric LA can still be used on spectral functions of this type, since they can be rewritten as a linear combination of three spectral functions for which Eq. 137 does hold [3]; that is,

$$\begin{aligned}\hat{g}(z) = \tfrac{1}{2}\{&\langle P_0^{1/2}(f_1 + \mathcal{T}_c f_2) | [z\mathbf{I} + \mathbf{A}_N]^{-1} | P_0^{1/2}(\mathcal{T}_c f_1 + f_2) \rangle \\ &- \langle P_0^{1/2} \mathcal{T}_c f_2 | [z\mathbf{I} + \mathbf{A}_N]^{-1} | P_0^{1/2} f_2 \rangle \\ &- \langle P_0^{1/2} f_1 | [z\mathbf{I} + \mathbf{A}_N]^{-1} | P_0^{1/2} \mathcal{T}_c f_1 \rangle\}.\end{aligned} \tag{137a}$$

It is easy to verify that the preceding three spectral functions satisfy the symmetry requirement Eq. 137 and thus can be directly calculated with the complex symmetric LA. In fact if $f_1 = \mathcal{T}_c f_1$ and $f_2 = \mathcal{T}_c f_2$ (e.g. one may choose $f_1$ and $f_2$ from the basis functions $\phi_i$ obeying Eq. 130), then Eq. 137a becomes the sum of three autocorrelation functions, two of which are the autocorrelations of $f_1$ and $f_2$.

Alternatively, the biorthogonal LA can be used to directly calculate the spectral function as advocated by Wassam [111, 112]. This approach would be most useful if only the cross-correlation of $f_1(q, \lambda)$ with $f_2(q, \lambda)$ is desired. However, the autocorrelations of $f_1(q, \lambda)$ and $f_2(q, \lambda)$ are usually needed to properly interpret the cross-correlation function. In this spirit, it is also necessary to apply the biorthogonal LA three times to get all the required information. The drawback of using the biorthogonal LA is that it requires twice as many matrix–vector multiplications as the complex symmetric LA and its numerical properties are not well understood. Also, in the relatively

simple models studied in detail, it is found that the symmetrized Fokker–Planck operators have simpler matrix element structures. This is not necessarily the case for more complex problems. (A Lanczos algorithm for *real* non-symmetric matrices has recently been described [215], and it can be useful for typical Fokker–Planck operators [41]).

### 3. *Example: The Planar Rotator*

The Fokker–Planck operator for the angular position and velocity of a planar rotator in the presence of a potential provides a good illustration of the general reduction method since it is not too complicated but does include most of the possible types of terms that can arise in the general Fokker–Planck operator. This model has been used to gain insight into the problem of including inertial effects in the SLE in the slow-motional regime [15]. In addition, Stillman and Freed have used it as a prototypical Fokker–Planck operator in a study of the stochastic modeling of the non-Markovian many-body features of diffusing molecules [41] by a procedure involving the augmentation of the basis set of dynamical variables. The symmetrization and discretization of the simple planar rotator model is discussed here as an illustration of the practical use of the basic ideas of Sections VI.B and VI.B.1. Augmented Fokker–Planck operators for the planar rotator including several different forms of fluctuating torques are given in Section VI.C after the general discussion on augmented Fokker–Planck equations.

The Fokker–Planck equation for the planar rotator problem is

$$\frac{\partial P(\gamma,\dot\gamma,t)}{\partial t} = -\Gamma(\gamma,\dot\gamma)P(\gamma,\dot\gamma,t), \tag{138}$$

where $\gamma$ and $\dot\gamma$ are, respectively, the angular position and velocity with respect to a fixed laboratory frame. The Fokker–Planck operator itself is the sum of a term that is odd under motion reversal,

$$\Gamma_1(\gamma,\dot\gamma) = -\Gamma_1(\gamma,-\dot\gamma) \tag{139}$$

$$= \dot\gamma\frac{\partial}{\partial\gamma} + \frac{F(\gamma)}{I}\frac{\partial}{\partial\dot\gamma}, \tag{140}$$

and a term that is even under motion reversal,

$$\Gamma_2(\gamma,\dot\gamma) = \Gamma_2(\gamma,-\dot\gamma) \tag{141}$$

$$= \frac{\beta k_B T}{I}\frac{\partial}{\partial\dot\gamma}\left(\frac{\partial}{\partial\dot\gamma} + \frac{\dot\gamma}{Ik_BT}\right), \tag{142}$$



where $\beta$ is a phenomenological friction coefficient, $I$ is the moment of inertia of the rotator, $k_B$ is Boltzmann's constant, and $T$ is the absolute temperature. The force on the rotator $F(\gamma)$ is derived from the potential function $V(\gamma)$,

$$F(\gamma) = -\frac{\partial}{\partial\gamma}V(\gamma). \tag{143}$$

The simplest case to consider is where the restoring force $F(\gamma)$ is not present. In this case, the Fokker–Planck operator

$$\Gamma(\gamma,\dot\gamma) = \dot\gamma\frac{\partial}{\partial\gamma} + \frac{\beta k_BT}{I}\frac{\partial}{\partial\dot\gamma}\left(\frac{\partial}{\partial\dot\gamma} + \frac{\dot\gamma}{Ik_BT}\right) \tag{144}$$

together with the Maxwell–Boltzmann equilibrium distribution

$$P_0(\gamma,\dot\gamma) = (2\pi)^{-3/2}(k_BT/I)^{-1/2}\exp\frac{-I\dot\gamma^2}{2k_BT} \tag{145}$$

satisfy the conditions of detailed balance. Applying the symmetrizing transformation to $\Gamma$ gives

$$\tilde\Gamma(\gamma,\dot\gamma) = \dot\gamma\frac{\partial}{\partial\gamma} + \beta\left(\alpha^{-2}\frac{\partial^2}{\partial\dot\gamma^2} - \tfrac{1}{4}\alpha^2\dot\gamma^2 + \tfrac{1}{2}\right), \tag{146}$$

where $\alpha = \sqrt{I/k_BT}$. The symmetrized Fokker–Planck operator assumes a simpler form when written in terms of the variables $q_1 = \gamma$ and $q_2 = \alpha\dot\gamma$,

$$\tilde\Gamma(q_1,q_2) = \alpha^{-1}q_2\frac{\partial}{\partial q_1} + \beta\left(\frac{\partial^2}{\partial q_2^2} - \tfrac{1}{4}q_2^2 + \tfrac{1}{2}\right). \tag{147}$$

It should be noted that the second term in Eq. 147, which arises from the irreversible part $\Gamma$, has the form of the differential equation for the parabolic cylinder functions [129, 130], or, to within an additive constant, a harmonic oscillator Hamiltonian. This observation, together with the particularly simple dependence on $\gamma$, suggest that the direct-product basis functions of complex exponentials and harmonic oscillator wavefunctions,

$$\psi_{lk}(q_1,q_2) = N_k e^{-(1/4)q_2^2 + ilq_1}H_k(q_2/\sqrt{2}), \tag{148}$$

where $k = 0, 1, 2, \ldots, K$, $l = 0, \pm1, \pm2, \ldots, \pm L$, $N_k = (2\pi)^{-3/4}(2^k k!)^{-1/2}$, and $H_k(q_2)$ are the Hermite polynomials (57, 129), will be well suited for this

problem. This basis is orthonormal but not complete since $K$ and $L$ are finite integers, and it is not invariant under the time reversal operation (see Eq. 130). The matrix elements of $\tilde{\Gamma}$ in this basis are

$$\langle \psi_{lk}|\tilde{\Gamma}|\psi_{mn}\rangle = \int_0^{2\pi}\int_{-\infty}^{\infty} dq_1\, dq_2\, \psi_{lk}^*(q_1,q_2)\tilde{\Gamma}(q_1,q_2)\psi_{mn}(q_1,q_2) \tag{149}$$

$$= \beta n\delta_{k,m}\delta_{l,n} + im\alpha^{-1}\delta_{k,m}[\sqrt{n+1}\,\delta_{l,n+1} + \sqrt{n}\delta_{l,n-1}]. \tag{150}$$

It is easy to verify that the matrix of $\tilde{\Gamma}$ is complex symmetric in this basis, that is, $\langle \psi_{lk}|\tilde{\Gamma}|\psi_{mn}\rangle = \langle \psi_{mn}|\tilde{\Gamma}|\psi_{lk}\rangle$.

In the presence of a potential [$F(\gamma) \neq 0$ cf. Eqs. 140 and 143], the matrix of $\tilde{\Gamma}$ is not necessarily complex symmetric in the basis specified by Eq. 148. A new set of basis functions that satisfy Eq. 130 and span the same subspace can be constructed by taking linear combinations of the old basis functions,

$$\phi_{mnp}(q_1,q_2) = (i)^m N_m([1+\delta_{n,0}])^{-1/2} e^{-(1/4)q_2^2} H_m(q_2/\sqrt{2}) \times \begin{cases} \cos(nq_1) & \text{if } p=1, \\ \sin(nq_1) & \text{if } p=-1, \end{cases} \tag{151}$$

where $n = 1, 2, \ldots, N$ if $p = -1$ and $n = 0, 1, 2, \ldots, N$ if $p = 1$. Also, $m = 0, 1, 2, \ldots, M$. The symmetry of the resulting matrix can be verified directly from the integral definition of the matrix elements by using the fact that $F(\gamma)$ is a real function and from the structure of the matrix elements of the position and momentum operators in the harmonic oscillator eigenfunction basis.

This basis is not the only possible choice that satisfies the symmetry condition given in Eq. 130. For instance, the basis derived from $\phi_{mnp}(q_1,q_2)$ by omitting the factor $(i)^m$ for even $m$ is also a valid basis that satisfies the symmetry requirement. In addition, any basis that can be expressed in terms of real linear combinations of the basis $\phi(q_1,q_2)$, (i.e. $A\phi_{mnp} + B\phi_{m'n'p'}$ with $A$ and $B$ real constants) will also satisfy the symmetry requirement. The choice of basis can be adapted to the calculation of a specific spectral function or class of spectral functions by taking advantage of the spatial symmetry group under which $\tilde{\Gamma}$ is invariant. The use of spatial symmetry adapted basis functions may reduce the dimension of the matrix.

We can use the functions given by Eq. 151 to illustrate the correlation functions for which Eqs. 137 and 137a hold, so that auto and cross correlations are readily obtained. That is, we may set $f_1 \propto e^{+(1/4)q_2^2}\phi_{mnp}$ and $f_2 \propto e^{+(1/4)q_2^2}\phi_{m'n'p'}$. In fact, any desired correlation function would be obtained directly or as a linear combination of such correlation functions.



### C. Extensions of the Fokker–Plank Approach

Stillman and Freed [41] have outlined an extension of the traditional Fokker–Planck approach that can be used to model the non-Markovian many-body features of diffusing molecules. It introduces, in a transparent manner, the basic physics of the relevant degrees of freedom and their couplings and is not restricted to linear transport laws. This approach can also describe both equilibrium and nonequilibrium dynamics but does require the independent specification of the proper equilibrium or stationary probability distribution for the system by independent means.

In this method, the set of relevant dynamical variables (see sections VI.A and VI.B.1) is augmented with stochastic bath variables assumed to have simple Markovian behavior. The augmented set of variables then represents a multidimensional Markov process that obeys a classical SLE. In general, the SLE does not obey the detailed balance conditions given previously since it ignores the back reaction of the dynamical variables on the bath variables. This is a well-known defect in the SLE approach (see section VII). To proceed, the back reaction(s) are incorported into the model by adding term(s) to the SLE to satisfy the detailed balance conditions. The resulting augmented Fokker–Planck equation (AFPE) describes the relaxation of the system to its stationary state, and under appropriate conditions, it also reduces to a classical Fokker–Plank equation for the initial set of dynamical variables. Augmented Langevin equations (ALE) that automatically satisfy the fluctuation–dissipation relationships may be readily obtained from the AFPE.

The relevant dynamical variables $\Delta$ of the system are assumed to obey an equation of motion of the form

$$\frac{d\Delta}{dt} = F(\Delta; \Xi(t), \lambda), \tag{152}$$

where $\Xi(t)$ represents the set of independent stochastic bath variables. Furthermore, the stochastic process for the bath variables is assumed to be stationary and Markovian with an associated master equation,

$$\frac{\partial P(\Xi,t;\lambda)}{\partial t} = -\Gamma_\Xi P(\Xi,t;\lambda). \tag{153}$$

With these assumptions, the SLE for the joint probability distribution function of the augmented set of variables can be written as [131–133]

$$\frac{\partial P(\Delta,\Xi,t;\lambda)}{\partial t} = -[\nabla_\Delta \cdot F(\Delta;\Xi,\lambda) + \Gamma_\Xi]P(\Delta,\Xi,t;\lambda). \tag{154}$$

Here the symbol $\nabla_\Delta$ represents the divergence with respect to the initial set of dynamical variables $\Delta$. The first term on the right side of Eq. 154 is then the Liouville form of Eq. 152. If the variables $\Delta$ and $\Xi$ are merged to form a new set $q$, Eq. 154 has the form of a generic Fokker–Planck equation (see Eqs. 87 and 91),

$$\frac{\partial P(q,t;\lambda)}{\partial t} = -\Gamma(q,\lambda)P(q,t;\lambda). \tag{155}$$

It must be reemphasized that Eq. 154 is incomplete in the sense that the back reaction of the dynamical variables on the bath has been ignored. Therefore, the stationary solution to Eq. 154 will, in general, only yield the correct Boltzmann distribution in the limit of infinite temperature. Another way of stating this is that the joint probability density in the augmented set of variables does not relax to thermal equilibrium! Clearly this is not satisfactory for present purposes. In order to obtain the physically correct approach to the stationary state, additional terms that have been omitted from Eq. 154 must be included. A sufficient condition for the solution of the SLE to relax to the proper stationary state is that it obey the detailed balance conditions stated in Section VI.B. An AFPE that satisfies the requirements of detailed balance can be obtained by adding appropriate reversible and/or irreversible drift terms to Eq. 155. The correct (or, at least reasonable) form of the terms must be determined from physical considerations. The same reasoning that went into specifying the proper stationary state is also applicable in determining these correction terms.

If desired, it is possible to generate a set of ALE from the AFPE [116]. The ALE for each variable $q_l$ in the augmented set can be written as

$$\frac{dq_l}{dt} = k_l(q,\lambda) + \sum_j g_{lj}(q,\lambda)x_j(t),$$

where the $x_j(t)$ are independent Gaussian $\delta$-correlated random functions of time with zero mean, that is,

$$\langle x_j(t)\rangle = 0,$$
$$\langle x_l(t+\tau)x_j(t)\rangle = \delta_{lj}\delta(\tau).$$

The functions $k_l(q,\lambda)$ and $g_{lj}(q,\lambda)$ are related to the drift and diffusion coefficients by

$$K_l(q,\lambda) = k_l(q,\lambda) + \frac{1}{2}\sum_{k,j}\frac{\partial g_{lj}(q,\lambda)}{\partial q_k}g_{kj}(q,\lambda),$$
$$K_{jl}(q,\lambda) = \sum_i g_{ji}(q,\lambda)g_{il}(q,\lambda).$$



The inversion of these equations to obtain the $g_{lj}(q,\lambda)$ and the $k_l(q,\lambda)$ is discussed in [116]. It is not, in general, unique in the absence of additional constraints [119].

An example of this scheme is the explicit introduction of fluctuating torques into the planar rotator problem. If $N(\gamma)$ and $T(\gamma,\dot\gamma,t)$ are the mean field torque and fluctuating torques, respectively, the SLE for the process is

$$\frac{\partial P(\gamma,\dot\gamma,t)}{\partial t} = -\left\{\dot\gamma\frac{\partial}{\partial\gamma} + I^{-1}\frac{\partial}{\partial\dot\gamma}[N(\gamma)+T(\gamma,\dot\gamma,t)]\right\}P(\gamma,\dot\gamma,t). \tag{156}$$

One physically plausible model for the fluctuating torque is given by assuming that angular position that minimizes the torque is undergoing a simple diffusion process, but the functional form of the mean square torque is dependent only on the deviation from the minimum. That is, it is reasonable to assume that

$$T(\gamma,\dot\gamma,t) = V_0\sqrt{Ik_BT}f(\gamma-\phi(t)),$$

where $\phi(t)$ is the stochastic variable characterizing the fluctuations in the position of the minimum of the torque, and the associated master equation for a simple diffusion in $\phi$ is valid:

$$\frac{\partial P(\phi,t)}{\partial t} = -\tau_\phi^{-1}\frac{\partial^2 P(\phi,t)}{\partial t^2}.$$

In addition, the stationary state must be specified. A natural choice is to let the equilibrium state be a Maxwell–Boltzmann distribution with respect to the angular velocity and assume that the mean torque is derivable from a potential function that depends only on $\gamma-\phi(t)$, that is, $N(\gamma) = -dU_N(\gamma-\phi)/d\gamma$ so that the equilibrium state is given by

$$\Phi(\gamma,\dot\gamma) = \frac{I\dot\gamma^2}{2k_BT} + \frac{U_N(\gamma)}{k_BT}.$$

With these assumptions about the nature of the equilibrium state and the form of the fluctuating torque, Eq. 156 does not satisfy the requirements for detailed balance. In particular, the divergence of the reversible probability current is nonzero (i.e., Eq. 111 is violated). The effects of the back-reaction terms can be included through the addition of a reversible drift term to Eq. 156 of the form $J_\phi = \sqrt{I/k_BT}V_0\dot\gamma g(\gamma-\phi)$, where $f(\gamma-\phi) = -dg/d\gamma$ such that the detailed balance conditions are satisfied.

Another physically interesting choice for the equilibrium distribution is to assume that the system relaxes to the instantaneous value of the fluctuating potential,

$$\Phi(\gamma, \dot{\gamma}) = \frac{I\dot{\gamma}^2}{2k_BT} + \frac{U_N(\gamma)}{k_BT} + \sqrt{Ik_BT}V_0 g(\gamma - \phi).$$

As before, the detailed balance conditions are not met without the addition of further terms. In this case, an irreversible drift term can be added to offset the nonzero irreversible probability current (cf. Eq. 112). This choice of potential is appropriate for cases where the torques relax on a time scale long compared to the angular position such as in the slowly relaxing local structure model used in the interpretation of magnetic resonance spectra of spin probes in liquid crystals and model membranes [16, 134].

The methods developed previously to symmetrize standard Fokker-Planck operators and the use of the LA to calculate spectral functions and spectral densities applies equally well to AFPE. In addition, this method which relies on the construction of an equation of motion for an augmented set of variables (AFPE or ALE) rather than a generalized Langevin equation (GLE), might prove useful in the area of stochastic molecular dynamics calculations since it is easy to include nonlinear couplings and the coefficients of the ALE are time independent. The absence of the memory kernels characteristic of GLE is due to the fact that the set of relevant variables has been extended to include the effects of these interactions. This can be thought of as a redefinition of the projection operators used to define the memory functions or as imposing some physically relevant, nontrival structure on the bath.

### D. Reduction of the Stochastic Liouville Operator to Complex Symmetric Matrix Form

An analysis of the general problem of the reduction of the stochastic Liouville operator to a complex symmetric matrix form is more complicated than the reduction for the general Fokker–Planck operator discussed previously. The source of the added complication is the coupling of classical and quantum mechanical degrees of freedom.

From general considerations, Hwang and Freed [135] have shown that spin-dependent force and torque terms must appear in the stochastic Liouville operator. Such terms represent the back reaction of the spin system on the classical orientational, positional, or velocity degrees of freedom. For example the coupling of a spin system to the SLE for the planar rotator in Eq. 15 would require that the mean field torque $N(\gamma)$ be replaced by $N(\gamma) - (i/2)[d\mathscr{H}_s^+(\gamma)/d\gamma]$, where the new term is the spin-dependent torque due to the angular dependence of the semiclassical spin Hamiltonian $\mathscr{H}_s(\gamma)$, and the plus superscript indicates an anticommutator superoperator. It arises from applying the Poisson bracket with respect to $\gamma$ in the classical part of the Liouville equation. The anticommutator feature is required to maintain a Hermitian density matrix at all times. Thus, the classical probability distribution must also be replaced by the generalized spin density operator of Eq. 7 and Section VII. The existence of a spin-dependent force term had earlier been inferred by Pedersen and Freed [67] in considering the problem of spin-dependent reactive trajectories of interacting radical pairs. A spin-dependent torque term was also inferred by Vega and Fiat [136]. It is easy to show that within the high-temperature approximation, the equilibrium potential should include the term $\mathscr{H}_s(\Omega)/k_BT$, that is the spin force or spin torque acts as an additional effective potential energy term. It is necessary, but need not be sufficient, to guarantee the relaxation of the spin system to thermal equilibrium [68]. Monchick [137] has more recently obtained similar results, and Wassam and Freed [138, 139] have provided a detailed theory that deals with the inclusion of such terms. Both these more recent works employ the Wigner distribution function to pass from a fully quantum-mechanical treatment to the semiclassical limit.



The preceding considerations show that the spin density matrix of the SLE should relax to a canonical distribution in $\mathscr{H}_s(\Omega)$ as well as to the equilibrium distribution for the molecular degrees of freedom. As long as $|\mathscr{H}_s(\Omega)/k_BT| \ll 1$, it is not necessary to explicitly include these spin force and spin torque terms into the SLE. Instead, the term $\Gamma\rho(\Omega, t)$ can be approximated by $\Gamma_{HT}[\rho(\Omega, t) - \rho_0(\Omega, t)]$, where $\Gamma_{HT}$ is the high-temperature limit of $\Gamma$ that does not include spin force or spin torque terms. This form of the SLE will still be consistent with a suitably extended detailed balance criterion, and after the symmetrization of $\Gamma_{HT}$, one can again construct complex symmetric matrix representations of the stochastic Liouville operator. The high-temperature form of the relaxation term in the SLE had been inferred by Freed [8], who found that it was required in order to have the general density matrix theory of the SLE reduce to the correct linear response result in the limit of weak irradiation fields (cf. Section II). This correspondence further required $\rho_0(\Omega, t)$ to be the canonical density matrix associated with the *instantaneous* value of $\mathscr{H}_s(\Omega)$, *including* the effects of the irradiating field. This high-temperature limit of the SLE is the form used in the next section.

We have found that this approach does indeed lead to a complex symmetric form of the stochastic Liouville operator in a careful treatment of the high temperature limiting case involving a generalization of the work of Lax [124, 125]. The previous work by Lynden-Bell [61, 62] and Pyper [59, 60] on the symmetries of the SLE is also useful. It should also be mentioned that a

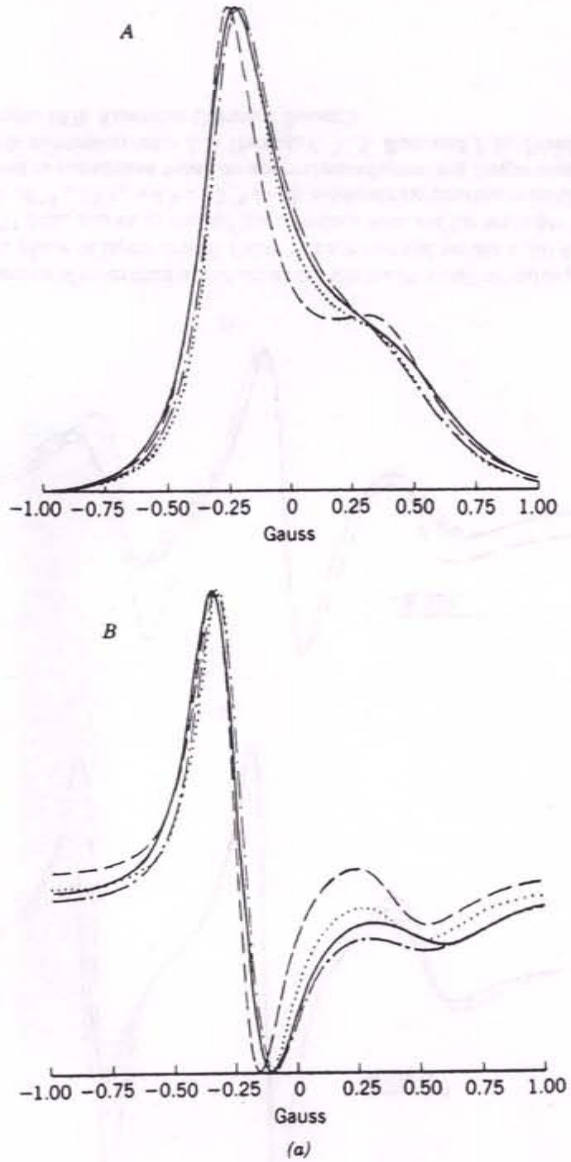



**Figure 11a.** Comparison of line shapes for axially symmetric $g$ tensor for different fre rotational diffusion models. (A) absorption lineshapes; (B) first-derivative lineshapes. Differen rotational diffusion models denoted by dotted lines for Brownian diffusion and solid lines fo motion described in full three-dimensional angular momentum space for Brownian particle witl damping coefficient $\beta = 4R$; dashed lines: motion described in one-dimensional angula momentum space for Brownian particle with $\beta = 4R$ and $R = 0.13|F|$; dotted-dashed lines: simpl free diffusion with $\beta/R = (R\tau)^{-1} = 4$. All have $\tau_R = 1.72 \times 10^{-7}$s. (From Ref. 15. Reprinted witl permission from G. V. Bruno and J. H. Freed, *J. Phys. Chem.* **78**, 935. Copyright 1974 Americar Chemical Society).

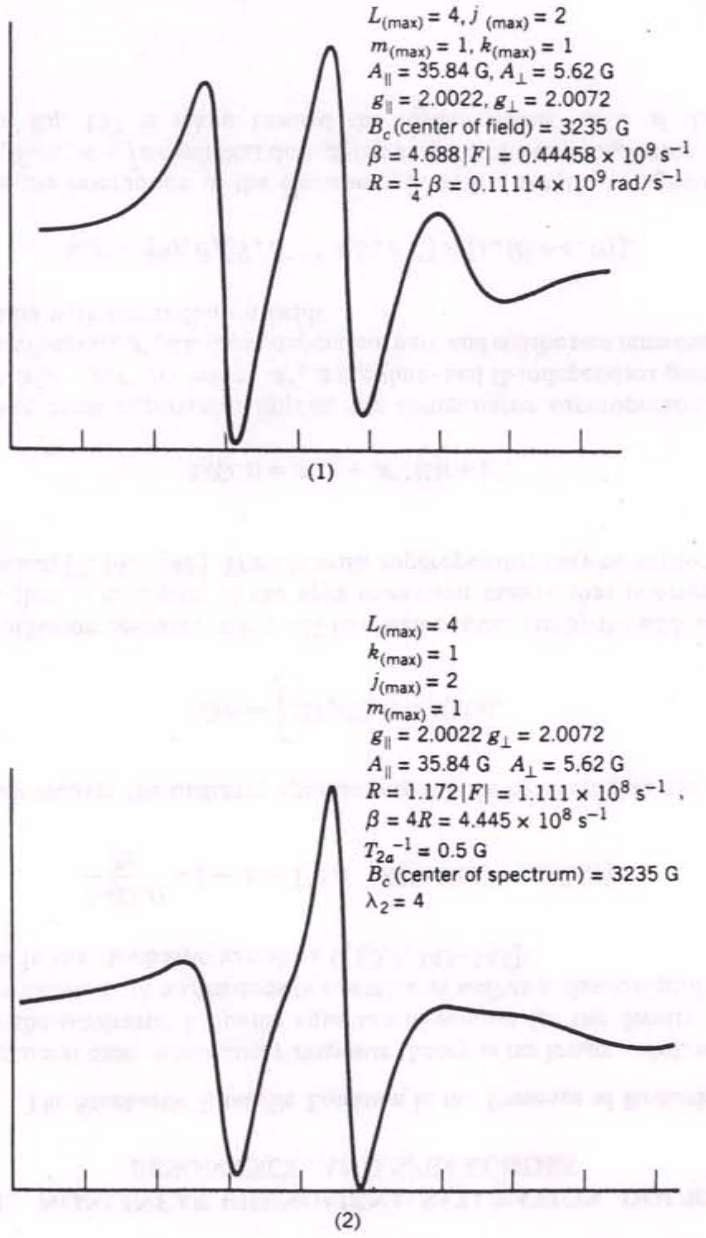



**Figure 11b.** Derivative lineshapes for axial nitroxide for motion described in full three-dimensional angular momentum space with $\beta = 4R$ and $\tau_R = 1.5 \times 10^{-9}$s. Case 1: No ordering potential. Case 2: Potential of $\lambda_0^2 = \frac{8}{3}$. (Approximate basis set utilized; cf. Ref. 142.) Separation between $x$-axis markers is 8.56G.

general prescription for symmetrizing the stochastic Liouville matrix for rotational diffusion problems was given previously [20].

In Figures 11*a* and 11*b* we show ESR spectral simulations [15, 140] based on the SLE for rotational diffusion, where $\Gamma$ is described by a Brownian Fokker–Planck equation in angular and angular momentum space [141]. Related results were obtained with an extended diffusion model [15, 140–142].

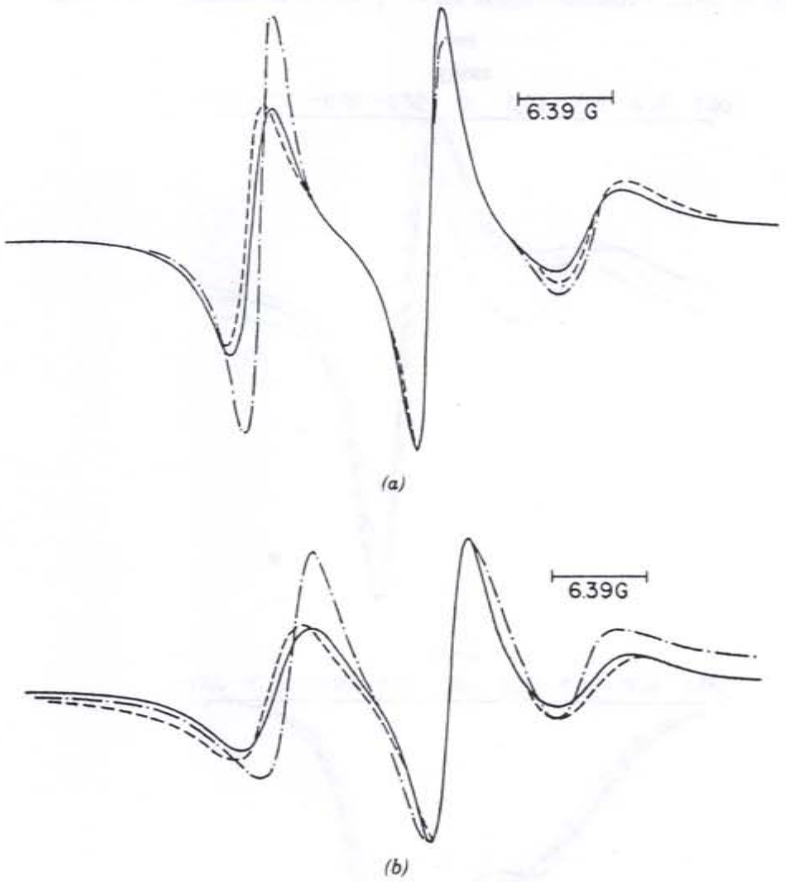



**Figure 12.** Comparison of experimental and simulated spectra for small nitroxide probe PD-TEMPONE in nematic phase of liquid crystal Phase V. Experimental results at (*a*) 45 °C, 3450 bars, and (*b*) 45 °C, 4031 bars, shown as dashed lines. Broken lines are for isotropic Brownian diffusion with (*a*) $2.25 \times 10^{-9}$ s; (*b*) $\tau_R = 4.3 \times 10^{-9}$ s with moderate asymmetric orienting potential; solid lines correspond to calculation based on approximate fluctuating torque model. (From Ref. 114. Reprinted with permission from J. S. Hwang, K. V. S. Rao, and J. H. Freed, *J. Phys. Chem.* **80**, 1790. Copyright 1976 American Chemical Society).



Inertial effects were shown to modify the spectrum for low friction. Methods equivalent to those of this section including the use of a symmetrized diffusion operator were employed [15]. Model calculations based on the fluctuating torque model (cf. Table 4) have also been performed [114] (cf. Figure 12) and are expected to be more relevant in viscous fluids than the inertial models.

## VII. NONLINEAR PHENOMENA: SATURATION, DOUBLE RESONANCE, AND SPIN ECHOES

### A. The Stochastic Liouville Equation in the Presence of Radiation

In the general case, when linear response theory is no longer valid, we may employ the stochastic Liouville equation of motion for the density matrix $\rho(\Omega, t)$, which is both a spin density operator as well as a classical probability function in the stochastic variables $\Omega$ [5, 8, 143–145]:

$$\frac{\partial \rho(\Omega, t)}{\partial t} = [-i\mathbf{L} - \Gamma(\Omega) - \mathbf{R}][\rho(\Omega, t) - \rho_0(\Omega, t)]. \tag{157}$$

One may recover the ordinary spin density matrix by averaging over $\Omega$:

$$\bar{\rho}(t) = \int d\Omega\, \rho(\Omega, t) = \overline{\rho(\Omega, t)}. \tag{158}$$

The diffusion operator in Eq. 157 has been augmented by the addition of a term **R** that is that part of the spin relaxation matrix that is orientation independent [7, 143–145]. The Liouville superoperator may be written as (cf. Eq. 20)

$$\mathbf{L}(\Omega, t) = \mathscr{H}_0^{\times} + \mathscr{H}_1^{\times}(\Omega) + \varepsilon^{\times},$$

where the cross superscript implies the commutator superoperator form, $\mathscr{H}^{\times}\rho \equiv \mathscr{H}\rho - \rho\mathscr{H}$. As before, $\mathscr{H}_0$ is the time- and $\Omega$-independent part of the spin Hamiltonian, $\mathscr{H}_1(\Omega)$ the $\Omega$-dependent part, and $\varepsilon(t)$ the new term due to the interaction with the radiation field:

$$\hbar\varepsilon(t) = \tfrac{1}{2}\hbar\gamma_e B_1[S_+ e^{-i\omega t} + S_- e^{i\omega t}] \equiv [\varepsilon_+(t) + \varepsilon_-(t)],$$

which is the interaction of the electron spin with a rotating magnetic field $\mathbf{B}_1 = B_1(\hat{i}\cos\omega t + \hat{j}\sin\omega t)$ that defines the $x$ axis of the rotating frame. Relaxation in Eq. 157 is taken toward the instantaneous value of the spin

Hamiltonian [8, 51, 98, 144, 145], that is,

$$\rho_0(\Omega,t)=P_0(\Omega)\frac{\exp[-\hbar\mathcal{H}(\Omega,t)/k_BT]}{\mathrm{Tr}\exp[-\hbar\mathcal{H}(\Omega,t)/k_BT]}\simeq P_0(\Omega)\frac{1}{N}\left[I-\frac{\hbar\mathcal{H}(\Omega,t)}{k_BT}\right], \tag{159}$$

where $N$ is the total number of spin eigenstates and $k_B$ is Boltzmann's constant. The approximate equality in Eq. 159 is a high-temperature approximation.

At this point it is useful to introduce the operator

$$\chi(\Omega,t)\equiv\rho(\Omega,t)-\rho_0(\Omega,t) \tag{160}$$

to simplify the form of the following equations. It follows from Eqs. 157 and 160 that

$$\frac{\partial\chi(\Omega,t)}{\partial t}+\frac{\partial\rho_0(\Omega,t)}{\partial t}=[-i\mathbf{L}-\mathbf{\Gamma}]\chi(\Omega,t) \tag{161}$$

(where **R** in Eq. 157 has been omitted for convenience) and from the high-temperature approximate form of Eq. 159:

$$\frac{\partial\rho_0(\Omega,t)}{\partial t}=\frac{i\hbar\omega}{Nk_BT}P_0(\Omega)[\varepsilon_+(t)-\varepsilon_-(t)]. \tag{162}$$

Except for the driving term involving the time derivative of $\rho_0(\Omega,t)$, Eq. 161 is formally similar to Eq. 10 for the spin operator in the Heisenberg representation. By introducing the symmetrized diffusion operator given by Eq. 17 in Section II, Eq. 161 can be converted to

$$\frac{\partial\tilde{\chi}(\Omega,t)}{\partial t}+P_0^{-1/2}(\Omega)\frac{\partial\rho_0(\Omega,t)}{\partial t}=[-i\mathbf{L}-\tilde{\mathbf{\Gamma}}]\tilde{\chi}(\Omega,t), \tag{163}$$

with

$$\tilde{\chi}(\Omega,t)=P_0^{-1/2}(\Omega)\chi(\Omega,t).$$

Thus, from Eq. 158, one has

$$\bar{\chi}(t)=\int d\Omega\,P_0^{1/2}(\Omega)\tilde{\chi}(\Omega,t)=\overline{P_0^{1/2}(\Omega)\tilde{\chi}(\Omega,t)}.$$

### B. General Methods of Solution

The general solution of Eq. 163 is obtained by first expanding $\tilde{\chi}(\Omega,t)$ in a Fourier series in the harmonics of the monochromatic radiation field,



$$\tilde{\chi}(\Omega,t)=\sum_{n=-\infty}^{\infty}\exp(in\omega t)\tilde{\chi}^{(n)}(\Omega,t). \tag{164}$$

(This may be regarded as a simple application of Floquet theory [129].) The Hermitian nature of $\rho$ and $\rho_0$ leads to the relation for the matrix elements of $\tilde{\chi}^{(n)}$,

$$\langle a|\tilde{\chi}^{(n)}|b\rangle=\langle b|\tilde{\chi}^{(-n)}|a\rangle^*. \tag{165}$$

In a CW experiment where the steady state is maintained by application of the radiation field at all times, one has

$$\tilde{\chi}^{(n)}(\Omega,t)=\tilde{\chi}_{ss}^{(n)}(\Omega).$$

That is, the Fourier coefficients, which are still spin density operators, are time-independent matrix elements of an ordinary spin operator.

The notation in common use in this field for various types of matrix elements of an ordinary spin operator $O$ is

$$O_\lambda\equiv O_{a^-,b^+}, \tag{166}$$

$$O_{\lambda\pm}\equiv O_{a^\pm,b^\pm}, \tag{167}$$

$$O_{\lambda_a^\pm,\lambda_b^\pm}\equiv O_{a^\pm,b^\pm}. \tag{168}$$

In eq. 166 the labeling is such that $\lambda$ refers to the $\lambda$th ESR transition between the states $a^-$ and $b^+$ where the lowercase letters indicate nuclear spin states and the minus and plus refer to the electron spin states $m_s=-\frac{1}{2}$ and $m_s=\frac{1}{2}$. If $a=b$, this is an allowed ESR transition; otherwise, it is a forbidden ESR transition. Equation 167 refers to the diagonal matrix elements, whereas Eq. 168 for $a\neq b$ is a "pseudodiagonal" matrix element [12, 97, 143–148] since it is diagonal with respect to the electron spin but the nuclear spin states are different. These are more properly included with the diagonal matrix elements (in part because of the very small differences in energy between nuclear spin states), and this will be done in what follows. In fact, the summations over $\lambda^\pm$ utilized in the following will imply both types of matrix elements Eqs. 167 and 168.

More generally, it is possible to expand an orientation-dependent operator $O(\Omega)$ in the direct product space of spin operators and the space of square integrable functions of $\Omega$ as was done in Eq. 25:

$$|O(\Omega,\omega)\rangle_o=\sum_{\lambda,m}o_{\lambda,m}(\omega)|\lambda,m\rangle,$$

$$|O(\Omega,\omega)\rangle_d=\sum_{\lambda^\pm,m}o_{\lambda^\pm,m}(\omega)|\lambda^\pm,m\rangle. \tag{169}$$

The subscripts $o$ and $d$ refer to the off-diagonal and diagonal subspaces. Then, utilizing Eqs. 162 and 164 for the off-diagonal spin matrix elements of Eq. 163,

$$\dot{z}^{(n)}_{\lambda,m}(t) = -i[n\omega - \omega_\lambda]z^{(n)}_{\lambda,m} - \sum_{m'}\langle\lambda,m|\tilde{\Gamma}|\lambda,m'\rangle z^{(n)}_{\lambda,m'}(t)$$
$$- i\sum_{\lambda',m'}\langle\lambda,m|\mathscr{H}^{\times}_1(\Omega)|\lambda',m'\rangle z^{(n)}_{\lambda',m'}(t)$$
$$- i\sum_{\lambda'^{\pm},m}\langle\lambda,m|\mathscr{H}^{\times}_1(\Omega)|\lambda'^{\pm},m\rangle\tilde{\chi}^{(n)}_{\lambda'^{\pm},m}(t)$$
$$- id_\lambda[\tilde{\chi}(n-1)_{\lambda^+,m} - \tilde{\chi}(n-1)_{\lambda^-,m}] + iq\omega d_\lambda\langle\lambda,m|P_0^{1/2}S_-\rangle. \quad (170)$$

The off-diagonal matrix elements of $\tilde{\chi}(t)$ in Eq. 170 are denoted $z^{(n)}_{\lambda,m}$, following the established convention [143–145]. Also, the other symbols in Eq. 170 are defined as

$$\hbar\omega_\lambda = E_{\lambda^+} - E_{\lambda^-}, \quad (171)$$

where $E_{\lambda^\pm} = (\hbar\mathscr{H}_0)_{\lambda^\pm,\lambda^\pm}$ is the zero-order energy of the $\lambda^\pm$ spin eigenstate of $\mathscr{H}_0$ and the "transition moment" for the $\lambda$th allowed or forbidden ESR transition $d_\lambda$ is given by

$$d_\lambda = \tfrac{1}{2}\omega_1(S_-)_{+,-} = \tfrac{1}{2}\omega_1, \quad (172)$$

with $\omega_1 = \gamma_e B_1$ and $q = \hbar/Nk_BT$. Note that in the last term on the right in Eq. 170 the matrix element $\langle\lambda,m|P_0^{1/2}S_-\rangle$ is nonzero only for allowed ESR transitions.

Similarly, the diagonal and pseudodiagonal matrix elements of Eq. 163 are

$$\dot{\tilde{\chi}}^{(n)}_{\lambda^\pm,m}(t) = -in\omega\tilde{\chi}^{(n)}_{\lambda^\pm,m} - \sum_{m'}\langle\lambda^\pm,m|\tilde{\Gamma}|\lambda^\pm,m'\rangle\tilde{\chi}^{(n)}_{\lambda^\pm,m'}(t)$$
$$- i\sum_{\sigma,m'}\langle\lambda^\pm,m|\mathscr{H}^{\times}_1(\Omega)|\sigma,m'\rangle\tilde{\chi}^{(n)}_{\sigma,m'}(t)$$
$$\pm 2d_\lambda[z^{(n+1)}_{\lambda^-}(t) - z^{(n+1)}_{\lambda^-}(t)], \quad (173)$$

where the subscript $\lambda^- \equiv \lambda$ refers to the $O_{a^-,b^+}$ matrix element as before (cf. Eq. 166), whereas $\lambda^-$ refers to the $O_{b^-,a^+}$ matrix element. In the third term on the right in Eq. 173 the summation index $\sigma$ runs over all $\lambda^-$, $\lambda^-$, and $\lambda^\pm$. Also note that from Eq. 165, $z^{(n-1)}_{\lambda^-}(t) = z^{(-n+1)*}_{\lambda^-}(t)$, whereas $\tilde{x}^{(n)}_{\lambda^\pm} = \tilde{\chi}^{(-n)*}_{\lambda^\pm}$.

The steady-state solutions of Eqs. 170 and 173 are obtained by setting $\dot{z}^{(n)}_{\lambda,m}(t) = \dot{\tilde{\chi}}^{(n)}_{\lambda^\pm,m} = 0$ to yield a set of time-independent coupled algebraic equations.

One sees, from Eqs. 170 and 173, that it is only through the effects of the radiation field, where the strength of the interaction with the spins is characterized by $\omega_1$, that the harmonics of the off-diagonal matrix elements $z^{(n+1)}_{\lambda^-,m'}$ and $z^{(n-1)}_{\lambda^-,m'}$ are coupled to the harmonics of the diagonal and pseudodiagonal elements $\tilde{\chi}^{(n)}_{\lambda^\pm,m}$. An analysis of these equations leads to the result that the extent of coupling depends essentially on the ratio $\omega_1/\omega_\lambda$, which is very small in the presence of large static magnetic fields. Thus, in this case it will be possible to decouple the various harmonics. Next, it will be shown that for high-field saturation cases, it is sufficient to retain only the $z^{(1)}_{\lambda^-,m'}$, $z^{(-1)}_{\lambda^-,m'}$ and the $\tilde{\chi}^{(0)}_{\lambda^\pm,m}$ terms. The higher harmonics become important in a variety of multiple-resonance schemes [98, 99, 143–145, 149] or experiments done at low static magnetic fields [5, 6].



Consider the power absorbed in a steady-state spectrum given by [51]

$$\mathscr{P} = \omega B_1\tilde{M}_y = -\tfrac{1}{2}i\omega B_1(M_+e^{-i\omega t} - M_-e^{i\omega t}), \quad (174)$$

where $\tilde{M}_y$ is the magnetization along the rotating $y$ axis, and the associated operators $\mathscr{M}_\pm$ are given in terms of the electron spin raising and lowering operators $S_\pm$ by

$$\mathscr{M}_\pm(t) = \mathscr{N}\hbar\gamma_e\,\mathrm{Tr}\,[\rho(\Omega,t)S_\pm] \quad (175)$$

(cf. Eqs. 8 and 9). By analogy to Eqs. 15 and 16, this can be rewritten as

$$\mathrm{Tr}_s[\bar{\rho}(t)S_\pm] = \mathrm{Tr}_s\overline{\rho(\Omega,t)S_\pm}. \quad (176)$$

Only the terms in $\rho(\Omega,t)$ that contribute to the net time-averaged power absorption via Eqs. 174–176 need be retained. From Eq. 160 we must consider the time evolution of both $\chi(\Omega,t)$ and $\rho_0(\Omega,t)$. We first note that $\rho_0(\Omega,t)$ given by Eq. 159 cannot contribute to the time-averaged absorption since (i) the time-independent terms in $\mathscr{H}_0 + \mathscr{H}_1$ have no components oscillating at $e^{i\omega t}$ needed to cancel the oscillations in Eq. 174, and (ii) the term in $\varepsilon(t)$ that has the needed oscillatory part is found to make contributions to the $M_\pm$ terms in Eq. 174 that are equal in magnitude and opposite in sign. The only terms in the expansion of $\tilde{\chi}(\Omega,t)$ that contribute to the net time-averaged power absorption are the $z^{(1)}_{\lambda^-,m}$ and $z^{(1)}_{\lambda^-,m}$ occurring in the form $z^{(1)''}_{\lambda,m} \equiv \mathrm{Im}\{z^{(1)}_{\lambda,m}\} = (2i)^{-1}[z^{(1)}_{\lambda^-,m} - z^{(-1)}_{\lambda^-,m}]$. The definition $z^{(1)'}_{\lambda,m} \equiv \mathrm{Re}\{z^{(1)}_{\lambda,m}\}$ will also be needed. Thus, from Eq. 174 the steady-state power absorption is given by

$$\mathscr{P} = 2\mathscr{N}\hbar\omega\sum_{\lambda,m} d_\lambda z^{(1)''}_{\lambda,m}\langle P_0S_-|\lambda,m\rangle. \quad (177)$$

In the sum over $\lambda$ only the allowed transitions contribute. In large static

magnetic fields only the coupling to $\tilde{\chi}^{(0)}_{\lambda^{\pm},m}$ is required, and these are simply interpreted in terms of deviations from the equilibrium population.

The structure of the coupled differential equations that emerges in the high-field limit from Eqs. 170 and 173 may be written in block matrix form as

$$\begin{pmatrix} \dot{\mathbf{z}}^{(1)}(t) \\ \dot{\mathbf{z}}^{(1)*}(t) \\ \dot{\tilde{\chi}}^{(0)}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{R}-i\mathbf{K} & \mathbf{0} & i\mathbf{d} \\ \mathbf{0} & \mathbf{R}+i\mathbf{K} & -i\mathbf{d} \\ i\mathbf{d}^{tr} & -i\mathbf{d}^{tr} & \mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{z}^{(1)}(t) \\ \mathbf{z}^{(1)*}(t) \\ \tilde{\chi}^{(0)}(t) \end{pmatrix} + \begin{pmatrix} i\mathbf{Q} \\ -i\mathbf{Q} \\ \mathbf{0} \end{pmatrix}. \quad (178)$$

The vector $\mathbf{z}^{(1)}(t)$ is a vector defined in an $(m \times \lambda)$-dimensional subspace with elements $z^{(1)}_{\lambda,m}$, where $\tilde{\chi}^{(0)}(t)$ is a vector defined in the $(\lambda^{\pm} \times m)$-dimensional subspace with elements $\tilde{\chi}^{(0)}_{\lambda^{\pm},m}(t)$, and $\mathbf{z}^{(1)*}(t)$ is the vector whose elements are the complex conjugates of the elements of $z^{(1)}(t)$.

The relaxation matrix $\mathbf{R}$ in Eq. 178 is the matrix representation of $\Gamma + \mathbf{R}$ from Eq. 157, and the coherence matrix $\mathbf{K}$ is the matrix representation of $\mathbf{L}$ from Eq. 170. By analogy with Section II, let $\mathbf{A}' = \mathbf{R} + i\mathbf{K}$ be the submatrix that governs the dynamics of the off-diagonal density matrix elements, noting, however, that $\mathbf{K}$ can include the effects of multiple quantum coherences, if present [144]. The symmetric matrix $\mathbf{W}$ plays the same role as $\mathbf{A}'$ but for the diagonal and pseudodiagonal density matrix elements [143–146]. The matrix $\mathbf{d}$ is the matrix of transition moments $d_\lambda$ that couples the space of electron spin transitions that contains the vector $z^{(1)}(t)$ to the space of populations and nuclear spin transitions that contains the vector $\chi^{(0)}(t)$, and it is not, in general, square. It has nonzero matrix elements only between electron spin transitions and their associated components in the subspace of diagonal and pseudo-diagonal matrix elements. The vector $\mathbf{Q}$ represents the driving terms in the space of electron spin transitions; its elements are given by the last term on the right side of Eq. 170.

The form of Eq. 178 is sufficiently general to be applicable to multiple-resonance schemes, and the methods for constructing the various matrices are discussed elsewhere [84, 143, 146]. It would, however, be necessary to include higher harmonics than $\mathbf{z}^{(1)}(t)$ and $\chi^{(0)}(t)$. Also note that the terms involving $z^{(0)}_{\lambda,m}$ and $\chi^{(1)}_{\lambda^{\pm},m}$ in Eq. 178 have been neglected. The omission of these terms amounts to neglecting the nonsecular terms from $\mathscr{H}^{\times}_1$; this is readily justified for slow motions in high field [5, 6].

Note that the form of Eq. 178 is of a *matrix* Bloch equation [143, 144], where $\mathbf{z}'$ and $\mathbf{z}''$ are the multidimensional analogues of $\tilde{M}_x$ and $\tilde{M}_y$, respectively, where $\chi$ is the analogue of $M_z - M_0$. In the limit of a very weak radiation field where $d_\lambda \to 0$, the couplings due to the $\mathbf{d}$ matrices can be neglected in Eq. 178. The neglect of these terms in the weak perturbation limit decouples the off-diagonal space from the diagonal and pseudodiagonal terms.



This allows one to solve just for the latter to obtain the power absorbed (cf. Eq. 174). In fact, in a steady state one has, $\chi^{(0)}_{ss} = 0$, and

$$\mathscr{P} = \tfrac{1}{2}\mathscr{N}\hbar\omega^2 q\omega_1^2 \langle v|[\mathbf{A}'(\omega)]^{-1}|v\rangle \quad (179)$$

from Eqs. 170, 174, and 178 with $|v\rangle$ given by Eqs. 25 and 27. This is seen to be proportional to the expressions for the absorption given by Eqs. 1 and 26. The actual experimentally observed signal is proportional to $\mathscr{P}/\omega_1$ [51].

For finite-amplitude irradiating fields it is necessary to consider the full matrix in Eq. 178, which is of complex symmetric form. Thus, the complex symmetric Lanczos and/or conjugate gradients algorithms can in principle be applied to Eq. 178 also.

### C. Steady-State Saturation and Double Resonance

Because of the complexity of Eq. 178, it is beneficial to consider some specific cases where simplifications are possible. For CW saturation [5, 97, 144, 150] and double-resonance schemes such as electron–electron double resonance (ELDOR) [98, 143, 144, 145, 149, 151] or electron–nuclear double resonance (ENDOR) [98, 143, 144], the steady-state solution of Eq. 178 involves setting $\dot{\mathbf{z}}(t) = \dot{\mathbf{z}}^*(t) = \chi(t) = 0$, so the spectrum may be solved either by tridiagonalizing the matrix in Eq. 178 using the LA or by solving the linear algebraic equations directly using the CG algorithm. The latter would appear to have the advantage because it is not possible to remove the sweep variable $\Delta\omega$ from all the diagonal elements. Thus, a Lanczos tridiagonalization would have to be performed for each value of $\Delta\omega$. However, it is possible to rearrange Eq. 178 in a partitioned form to obtain

$$\begin{pmatrix} \mathbf{R}-i\mathbf{K}-\mathbf{S} & \mathbf{S} \\ \mathbf{S} & \mathbf{R}+i\mathbf{K}-\mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{z} \\ \mathbf{z}^* \end{pmatrix} = \begin{pmatrix} -i\mathbf{Q} \\ i\mathbf{Q} \end{pmatrix}, \quad (180)$$

where the saturation matrix $\mathbf{S}$ is defined by

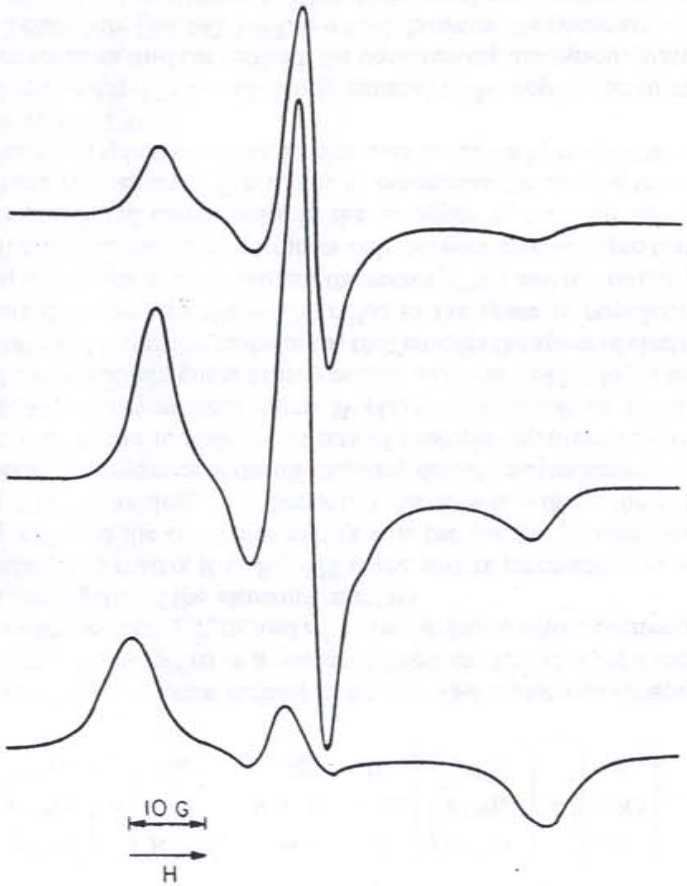
$$\mathbf{S} = \mathbf{d}\mathbf{W}^{-1}\mathbf{d}^{tr}. \quad (181)$$

Equation 180 may now be diagonalized only once to yield the entire spectrum. The details of constructing the symmetric $\mathbf{W}$ matrix and dealing with its singular nature are discussed elsewhere [5, 97, 144, 148]. The solution of Eq. 180 in previous studies has been accomplished by tri-diagonalization using Rutishauser's variant of Givens method (cf. Section IV) followed by diagonalization, but not by the LA. Thus, further work is needed to determine whether the solution of Eq. 180 using the LA or the solution of Eq. 178 (with time derivatives equal to zero) by the CG algorithm is preferred. Methods

based on Gaussian elimination (see Section IV) have previously been used to analyze cases involving modulation of the static magnetic field with detection of various harmonics [98, 99].

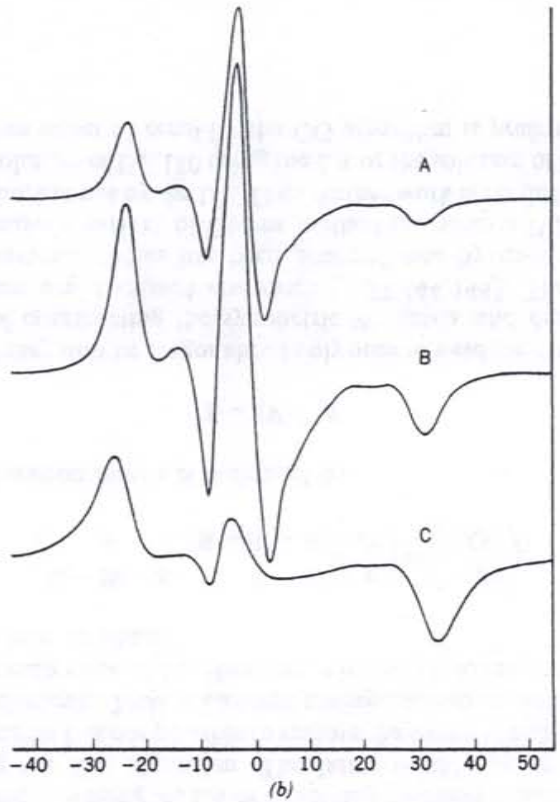
In Figure 13 we show results of an experimental study on ESR saturation in the slow-motional regime along with the associated calculations.

The general features of the method can be illustrated by considering the simple example of the ESR spectrum of a paramagnetic molecule with an axially symmetric $g$ tensor that is tumbling in an isotropic liquid. In the secular approximation, the orientation-dependent part of the spin Hamiltonian is




(a)






(b)

**Figure 13.** (*a*) Experimental saturation spectra for peroxylamine disulfonate in viscous solvent: "transition moments": (A) 0.025 G; (B) 0.079 G; (C) 0.45 G. (*b*) Simulated spectra for free diffusion with $\tau_R^{free} = 2.0 \times 10^{-8}$ s and appropriate transition moments. Magnetic tensors taken as axially symmetric for simplicity (From Ref. 97b).

simply

$$\mathscr{H}_1(\Omega) = \mathscr{F}\mathscr{D}^2_{00}(\Omega)S_z,$$

with

$$\mathscr{F} = \frac{2\beta_e B_0}{3\hbar}(g_\parallel - g_\perp),$$

where $g_\parallel$ and $g_\perp$ are, respectively, the parallel and perpendicular components of the $g$ tensor. It is possible to introduce the orientation-independent contribution to the linewidths, namely the relaxation rates $T_2'^{-1}$ and $T_1'^{-1}$, by

including these terms in the operator **R**. With the inclusion of these terms and the definition

$$\kappa_{L,L'} = [(2L+1)(2L'+1)]^{-1} \begin{pmatrix} L & 2 & L \\ 0 & 0 & 0 \end{pmatrix} \mathscr{F},$$

Eqs. 170 and 173 yields

$$[\Delta\omega - i(T_1'^{-1} + \tau_L^{-1})C_{00}^L(\Delta\omega) - \sum_{L'} \kappa_{L,L'} C_{00}^{L'}(\Delta\omega) + \sqrt{2}\, d b_{00}^L(\Delta\omega) = q\omega\, d\delta_{L,0}$$

and (182)

$$-i(T_1'^{-1} + \tau_L^{-1})b_{00}^L(\Delta\omega) + \sqrt{2}\, d\, \mathrm{Im}\{C_{00}^L(\Delta\omega)\} = 0, \quad (183)$$

where $\tau_L^{-1}$ are the eigenvalues of the isotropic diffusion operator given in Section II. In Eqs. 182 and 183 the standard notation [144] has been used where $z_{\lambda,m}^{(1)} \to C_{00}^L$ and $1/\sqrt{2}(\chi_{\lambda^-,m}^{(0)} - \chi_{\lambda^+,m}^{(0)}) \to b_{00}^L$. The matrix elements of the problem are

$$R_{L,L'} = (-T_2'^{-1} + \tau_L^{-1})\delta_{L,L'},$$
$$K_{L,L'} = \Delta\omega + \kappa_{L,L'},$$
$$\hat{W}_{L,L'} = (T_1'^{-1} + \tau_L^{-1})\delta_{L,L'},$$
$$Q_L = q\omega d\delta_{L,0}.$$

Note that **R** and $\hat{\mathbf{W}}$ are diagonal in this representation but **K** is not. In this problem $\hat{\mathbf{W}}$ has nonzero elements only in the subspace spanned by the linear combinations $b_{00}^L$. By the structure of the vector **Q**, the steady-state absorption lineshape is determined solely by $\mathrm{Im}\{C_{00}^0\}$. The physically interesting feature here may be appreciated by first realizing that $b_{00}^0$, which represents the ensemble-averaged deviation of the difference in spin population from its equilibrium value due to saturation effects, has a relaxation rate determined by $T_1'^{-1}$ alone. However, the $b_{00}^L$ for $L > 0$, which represent the nonspherically symmetric components of this deviation due to saturation, relax like $T_1'^{-1} + \tau_L^{-1}$. The rotational motion can relax the saturation at one region of the spectrum by transferring the saturated spins to another region of the spectrum.

Another example is the ELDOR signal of the same radical [149]. Here, two radiation fields, a pumping and an observing field, are employed. Therefore, one must label the expansion coefficients with two indices that keep track of the harmonics of each field. Equations 182 and 183 now become

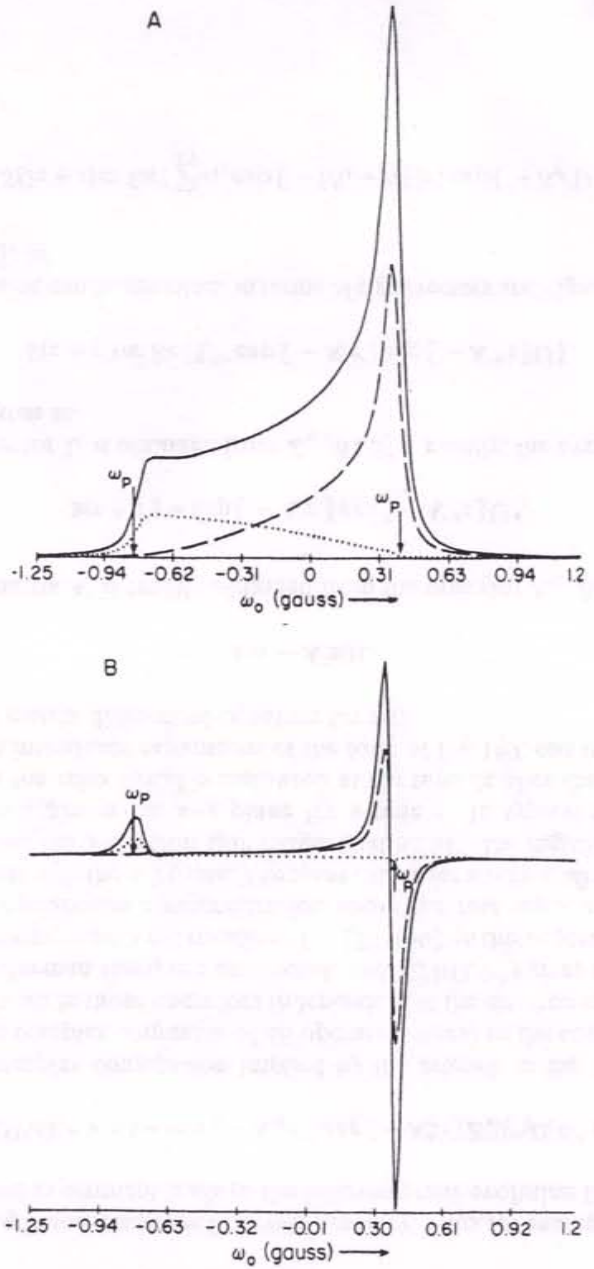



**Figure 14.** Predicted ELDOR lineshapes for axially symmetric $g$ tensor undergoing isotropic Brownian rotational diffusion. Pump frequency $\omega_p$ set at fixed value as shown and observing frequency $\omega_o$ is swept. Note how CW saturating effect of pump only partially transmitted to rest of spectrum by rotational motion. Here $\tau_R = 5.75 \times 10^{-6}$s. $T_1 = 5.7 \times 10^{-6}$s; (A) absorption; (B) first derivative. Solid line: pure ESR. Dashed and dotted lines: ESR line with pump on. (From Ref. 149.) Courtesy of North-Holland Publ. Co.

$$[\Delta\omega_\alpha - i(T_2'^{-1} + \tau_L^{-1})]C_{00}^L(\alpha) - \sum_{L'} \kappa_{L,L'} C_{00}^{L'}(\alpha) + \sqrt{2} d_\alpha b_{00}^L(\Delta\omega) = q\omega_\alpha d_\alpha \delta_{L,0}$$

and (184)

$$-i(T_1'^{-1} + \tau_L^{-1}) b_{00}^L(\Delta\omega) + \sqrt{2} \sum_\alpha d_\alpha \mathrm{Im}\{C_{00}^L(\alpha)\} = 0, \quad (185)$$

where $\alpha = o$ or $d = p$, referring to the observing and pumping fields, respectively. Here, $d_\alpha = \omega_{1,\alpha}/2$ and $C_{00}^L(o)$ and $C_{00}^L(p)$ are the $z_{\lambda,m}^{(n_o,n_p)}$ for the $n_o = 1$, $n_p = 0$, and $n_o = 0$ and $n_p = 1$ harmonics, respectively. By setting $\Delta\omega_o$ to be in resonance at positions in the absorption line different from $\Delta\omega_p$, it is possible to monitor the transfer of the saturating effects of the pumping field to the region around the observing field. This absorption of the observing field is proportional to $\mathrm{Im}\{C_{00}^L(o)\}$, which is a function of both $\Delta\omega_o$ and $\Delta\omega_p$. In this simple example, the "saturation transfer" is induced by the rotational diffusion modes with rates $\tau_L^{-1}$, as in the previous example. This "saturation transfer" effect in CW ELDOR is illustrated by the calculations shown in Figure 14.

### D. Spin Echoes and Two-Dimensional ESE

For general time domain experiments such as saturation recovery [143, 145, 152] or electron spin echoes [84], one must solve the complete form of Eq. 178 either in the time or frequency domain. However, there are special cases for which the saturation recovery problem can yield analytical forms [143, 144, 145]. Also, electron spin echoes and free induction decays may be greatly simplified by considering somewhat idealized models [26, 28, 30, 33, 35, 146]. Let us briefly consider the latter.

An idealized description of spin echoes is to assume that the radiation field is intense and of sufficiently short duration so that the sole effect of a pulse is a rotation of the density matrix by an appropriate angle. Thus, we assume that $\varepsilon(t)$ dominates the spin Hamiltonian in the rotating frame (i.e., for the equations written in terms of the $z_{\lambda,m}^{(1)}$) during the essentially infinitesimal time that the radiation field is on. Then, in the absence of any radiation, Eqs. 170 and 173 with $d_\lambda = 0$ give the (uncoupled!) time evolution of the off-diagonal and diagonal subspaces during such evolution periods. Thus, for the first evolution period we have

$$\rho_{rf}(\Omega, t + \tau) = \exp[-i\varepsilon^\times \tau]\rho_{rf}(\Omega, t), \quad (186)$$

where the subscript rf implies that only $z_{\lambda,m}^{(1)}$ and $\chi_{\lambda\pm,m}^{(0)}$ are considered in the expansion of $\rho(\Omega, t)$ according to Eqs. 160, 164, and 169. Also, $\tau$ is considered as an infinitesimal quantity. The result of this type of idealized pulse is simply a



rotation of the density matrix that only affects the electron spins. Whereas for the second period we have

$$\dot{\chi}_{rf}^{(n)}(\Omega, t) = -\mathbf{A}_{rf}^{(n)} \chi_{rf}(\Omega, t), \quad (187)$$

where

$$\mathbf{A}_{rf}^{(n)} \equiv -i[\mathscr{H}_0^\times - n\omega \hat{S}_z^\times + \mathscr{H}_1^\times(\Omega)] - \Gamma(\Omega)$$

is the rotating frame stochastic Loiuville operator. Thus, for example, a $\pi/2$–$\tau$–$\pi$–$\tau'$ spin echo experiment leads to the following time evolution for $\mathbf{Z}^{(1)}(\Omega, t)$:

$$\mathbf{Z}^{(1)}(\Omega, \tau + \tau') = \exp[-\mathbf{A}_{rf}\tau'] \exp[-\mathbf{A}_{rf}^*\tau] \mathbf{Z}_{\pi/2}^{(1)*}(\Omega, 0^+), \quad (188)$$

where the complex conjugation implied by the asterisk in Eq. 188 is more precisely the complex conjugate of all operators linear in the components of electron spin while those operators independent of the electron spin (such as the nuclear Zeeman term) are unaffected, and $\mathbf{Z}_{\pi/2}^{(1)}(\Omega, 0^+)$ gives the value of $\mathbf{Z}^{(1)}(\Omega, t)$ resulting from a $\pi/2$ rotation of $\rho_0$ [28, 146]. In this sequence, one first rotates the equilibrium $z$ magnetization about the rotating $x$ axis into the rotating $y$ axis with the $\pi/2$ pulse. The spins evolve for a time $\tau$, after which the $\pi$ pulse rotates the $y$ electron spin magnetization into the negative $y$ axis. It then evolves again in the $x$–$y$ plane for a time $\tau'$. In typical pulsed ESR experiments the echo signal is measured at the time $2\tau$ after the first pulse.

After one introduces expansions of the form of Eq. 169, one obtains from Eq. 187 the matrix differential equation for $\mathbf{z}(t)$:

$$\dot{\mathbf{z}} = -\mathbf{A}'\mathbf{z}(t),$$

where the matrix $\mathbf{A}'$ is readily obtained from the operator $\mathbf{A}_{rf}$. Also, Eq. 188 yields

$$\mathbf{z}(\tau + \tau') = \exp[-\mathbf{A}'\tau'] \exp[-\mathbf{A}'^*\tau]\mathbf{U}^*,$$

where the vector $\mathbf{U}$ is obtained from $\mathbf{Z}_{\pi/2}(\Omega, 0^+)$. Finally, the averaged signal may be written as

$$S(\tau + \tau') \propto \mathrm{Re}\{\mathbf{U}^{tr} \exp[-\mathbf{A}'\tau'] \exp[-\mathbf{A}'^*\tau]\mathbf{U}\}. \quad (189)$$

This expression can be rewritten in terms of eigenvectors and eigenvalues of $\mathbf{A}'$ (cf. Eq. [28]) as

$$S(2\tau + t) \propto \mathrm{Re}\{\sum_{i,j} c_{ij} \exp[-(\Lambda_i + \Lambda_j^*)\tau] \exp[-\Lambda_i t]\}, \quad (190)$$

where $\tau, t > 0$ and $t \equiv \tau' - \tau$. The coefficients $c_{ij}$ are given by

$$c_{ij} = \sum_{k,l,m} U_k O_{ki} O_{il} O^*_{ij} O^*_{mj} U^*_m, \tag{191}$$

where $\mathbf{O}$ is the complex orthogonal matrix that diagonalizes the matrix $\mathbf{A}'$. These eigenvectors are referred to as "dynamic spin packets" [30]. After performing a two-dimensional Fourier transform of Eq. 190 with respect to $2\tau$ and $t$ and using Eq. 192, we recover the general expression for two-dimensional ESE spectra for which Eq. 78 is a special case. The LA may easily be applied, as already discussed in Section V. Equation 191 shows that, in general, one would require the full $\mathbf{O}$ matrix to obtain the signal. This requirement could significantly reduce the power of the LA (but see what follows). However, in the very slow-motional regime where actual experiments of this type are performed, $|\mathscr{H}^x_1|/|\mathbf{\Gamma}| \gg 1$ so that $\mathbf{O}$ must be very nearly real. By setting $\mathbf{O}^* \simeq \mathbf{O}$ in Eq. 191, the approximate form of $c_{ij}$ is

$$c_{ij} \simeq (\mathbf{O}^{tr}\mathbf{U})^2_j \delta_{ij} = c^2_j \delta_{ij}. \tag{192}$$

The calculation of the approximate $c_{ij}$ from Eq. 192 is much simpler than accumulating the full transformation matrix $\mathbf{O}$. The approximate $c_{ij}$ are easily computed by storing only one vector containing the elements $(\mathbf{Q}^{tr}_D\mathbf{U})$, where the matrix $\mathbf{Q}_D$ is now the matrix that diagonalizes the Lanczos tridiagonal matrix. This procedure is implied in Eq. 78. However, the full matrix representation is needed in Eq. 191. This matter can be studied by first writing $\mathbf{O} \simeq \mathbf{Q}_L\mathbf{Q}_D$, where $\mathbf{Q}_L$ is the matrix of Lanczos vectors that reduces $\mathbf{A}'$ to tridiagonal form $(\mathbf{Q}^{tr}_L\mathbf{A}'\mathbf{Q}_L = \mathbf{T}_n)$. Now Eq. 191 may be approximated as

$$c_{ij} \simeq \sum_{i,r,s} Q_{L,1i} Q_{D,ir} Q_{L,rl} Q^*_{L,is} Q^*_{D,sj} Q^*_{D,1j}, \tag{193}$$

where we have used the fact that $\sum_k U_k Q_{L,kp} = \delta_{p,1}$, since the starting vector $\mathbf{U}$ is just the first Lanczos vector. [Note that the very slow-motional form in Eq. 192 can further be approximated as $c_{ij} \simeq (Q_{D,j1})^2 \delta_{ij}$.] The sum over $i$ in Eq. 193 is over the $N$-dimensional space spanned by the original basis states, whereas the $r$ and $s$ indices refer to the $n_s$-dimensional subspace spanned by the Lanczos vectors. The needed Lanczos vectors are seen to be the ones required to adequately represent the relevant eigenvectors and eigenvalues in Eq. 190, and these are the same as those that are important in the approximate form Eq. 192. Thus, it seems reasonable to expect that the MTS required for the approximate solution, discussed in Section V.B and Table 2, would be adequate for the original $N$-dimensional basis set, whereas the $n_s$-dimensional

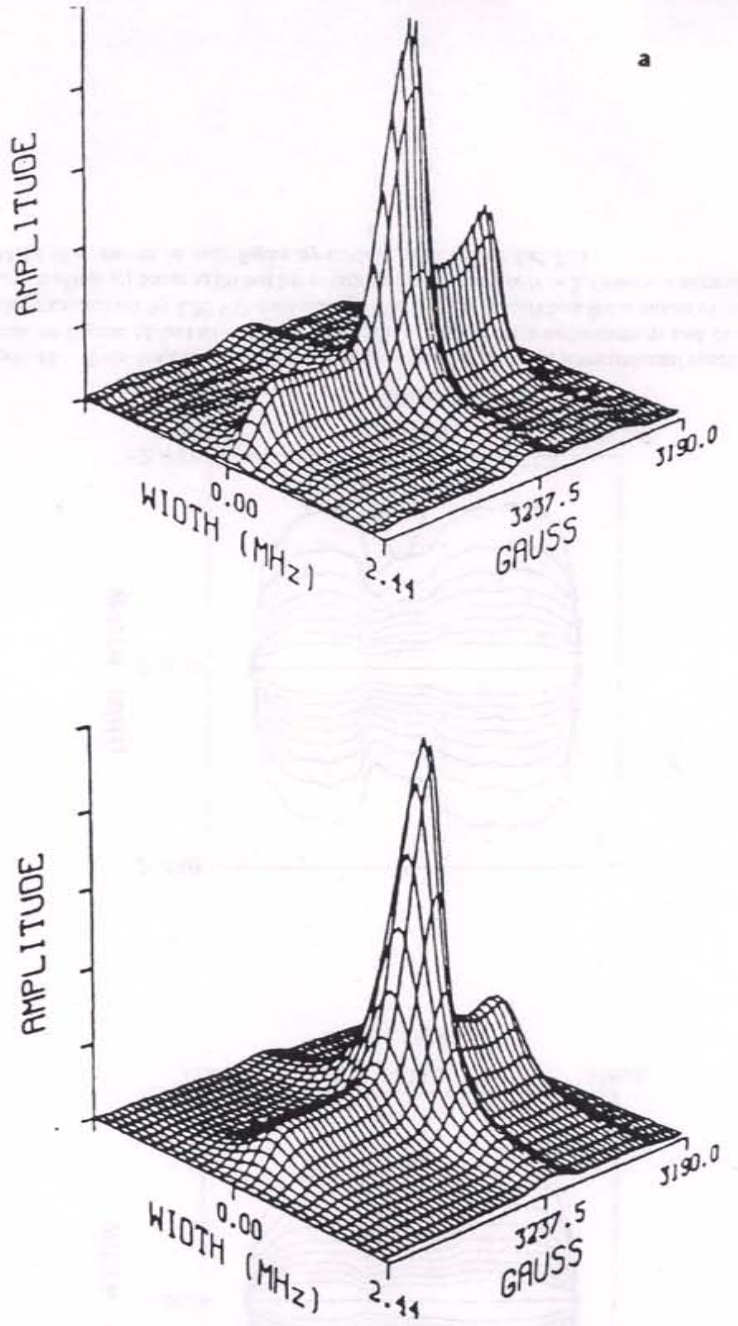
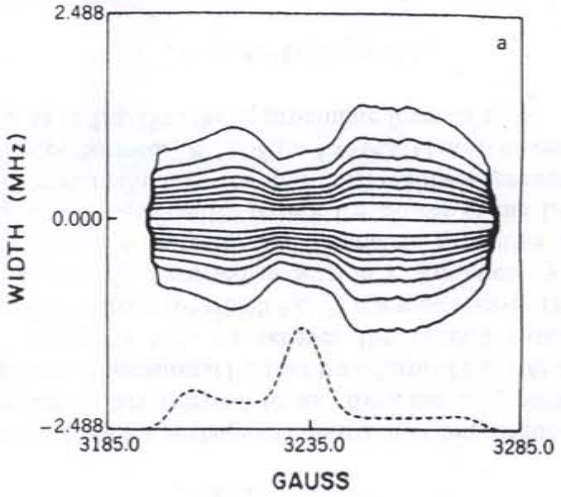


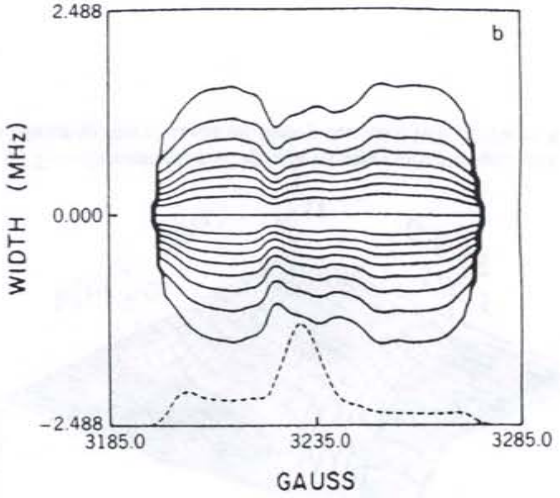

**Figure 15.** (*a*) Two-dimensional ESE spectra for small nitroxide probe in viscous solvent at −100 °C. (*b*) Simulated spectrum based on theory described in text. (From Ref. 30.)

set of Lanczos vectors from the approximate solution would suffice for the actual calculation of the two-dimensional ESE spectrum. Thus, the number of matrix elements in $\mathbf{Q}_L$ in Eq. 193 would be $N \times n_s$.

Rather than the time-consuming accumulation of the transformation matrices, it should be more efficient to adopt the strategy proposed by Cullum and Willoughby [89]. This would involve applying the LA to obtain $\mathbf{T}_n$ followed by the diagonalization of $\mathbf{T}_n$ by the $QR$ or $QL$ procedure without accumulating the transformation matrix $\mathbf{Q}_D$. After this is complete, the subset of eigenvectors of $\mathbf{T}_n$ corresponding to slowly decaying modes could be reconstructed by inverse iteration [80, 89] and the LA reapplied to $\mathbf{A}'$ to supply the Lanczos vectors required for the transformation of the subset of eigenvectors of $\mathbf{T}_n$ into a set of approximate eigenvectors of $\mathbf{A}'$. This procedure has the advantage that the Lanczos vectors need not be stored; they are regenerated when needed. Alternatively, if it appears that the reapplication of the LA is more time consuming than the storage of the Lanczos vectors, they can be stored. This might often be the case in the class of problems under consideration, where $n_s \ll N$. Another advantage of this approach is that, in general, the number of relevant eigenvectors, $n_e$, is typically $n_e \ll n_s$, so even less storage and fewer matrix multiplications are required to calculate the $c_{lj}$ by Eq. 193.

The theoretical curve in Figure 10 was obtained from Eq. 190 by setting $t = 0$ corresponding to the echo maximum and then stepping out $\tau$ to get the $T_2$-type decay. Figure 15 illustrates experimental and simulated two-dimensional ESE spectra for a nitroxide in a viscous fluid. The variation of



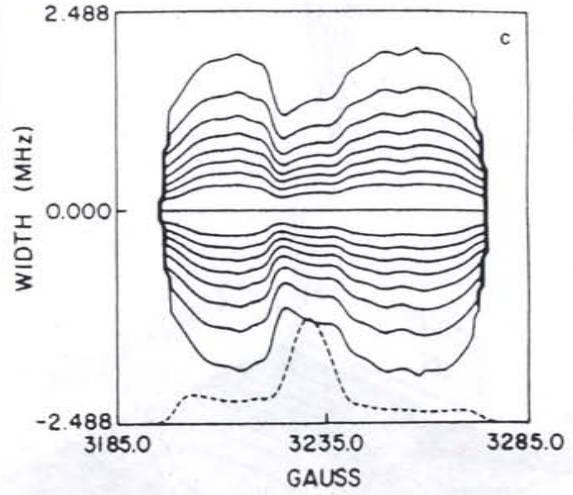




**Figure 16.** Two-dimensional ESE normalized contour plots. (*a*) Experimental spectrum for same case as Figure 15 but at −75 °C, with signal-to-noise ratio enhancement and dead-time corrections calculated by LPSVD methods. (*b*) Theoretical calculation for a model of isotropic Brownian motion. (*c*) Same as (*b*) but for rotational anisotropy of $N = 2$. Contours normalized to the 0-MHz slice, shown in each figure by dashed lines. (From Ref. 32.)

$T_2$ across the spectrum is illustrated in the experimental and theoretical contour plots shown in Figure 16. In the very slow-motional regime studied, the CWESR spectra are no longer sensitive to motion, but these two-dimensional ESE spectra remain very sensitive to the motion and to the microscopic model. Note that the two-dimensional spectral resolution and signal-to-noise ratio in the experimental data have been very significantly enhanced by modern techniques of data processing based on linear prediction [32, 153, 154].

Finally, we note that when the approximation of Eq. 192 is not valid and Eq. 191 must be used, the two-dimensional Fourier transform of Eq. 189 with respect to $\tau$ and $\tau'$ can lead to cross-correlations between the different dynamic spin packets. Analogous types of cross-correlations will show up in the next section.

### E. Stimulated Echoes, Magnetization Transfer, and Two-Dimensional Fourier Transform Spectroscopy

We now consider more sophisticated but still idealized pulse sequences. An important pulse sequence is the stimulated echo sequence $\pi/2$–$\tau_1$–$\pi/2$–T–$\pi/2$–$\tau_2$–echo. In this case, the second $\pi/2$ pulse rotates the $y$ magnetization into alignment along the negative $z$ axis. After evolving for a time $T$ along the $z$ axis, it is returned to the $x$–$y$ plane by the third $\pi/2$ pulse, and an echo is formed. By stepping out the time $T$, one can study relaxation of the $z$ magnetization. We refer to this as a $T_1$-type, or magnetization transfer, experiment.

This more sophisticated sequence yields the following expression for the desired stimulated echo signal [33, 146]:

$$S(T+\tau_1+\tau_2) \propto \mathrm{Re}\{\sum_p b_p(\tau_1,\tau_2)\exp[-T/\tau'_p|\}, \tag{194}$$

where $\tau'_p$ are the eigenvalues of the W matrix in Eq. 178 representing relaxation of the diagonal and pseudodiagonal density matrix elements. Also,

$$b_p(\tau_1,\tau_2) = \sum_{l,k,m,n,q,s} (U_q O_{o,qn}) O_{o,mn} O_{d,mp} O_{d,kp} O^*_{o,kj} (O_{o,sj} U_s)^*. \tag{195}$$

In Eq. 195 we distinguish the complex orthogonal transformation that diagonalizes A′ as in Eq. 191 by $\mathbf{O}_o$ and the complex orthogonal transformation that diagonalizes W by $\mathbf{O}_d$. One might expect W to be a real symmetric matrix. But because it couples the diagonal and pseudodiagonal matrix elements (which rigorously are off-diagonal density matrix elements), it is, in general, complex and symmetric. However, all complex eigenvalues and eigenvectors occur in complex-conjugate pairs, such that true diagonal density matrix elements, which represent populations, are always real [144–148].



Equation 195 is based on the simplifying assumption that the whole spectrum is irradiated by the idealized pulse. Also it assumes the existence of a "congruence" between the basis sets used to expand the off-diagonal and diagonal subspaces. This congruence, or one-to-one relationship, may be established via the **d** matrix of Eq. 178 in a number of useful cases [144–148]. In order to obtain the full time evolution of the signal, one must separately diagonalize **A** and **W**. Furthermore, there are no known simplifications analogous to Eq. 192 for the two-dimensional ESE case. Thus, in principle, one requires the full transformation matrices $\mathbf{O}_o$ and $\mathbf{O}_d$. As already noted in other contexts, this would greatly reduce the speed of the LA, and this procedure has not been used to date on this problem, although the Rutishauser algorithm has been employed [146].

Let us consider what would be involved in employing the LA in such calculation. We start by first comparing Eq. 195 for $b_p(\tau_1,\tau_2)$ versus Eq. 191 for $c_{ij}$. In the latter the complicating feature was the appearance of the sum $\sum_i O_{o,il} O^*_{o,ij}$, whereas in the former it is the double sum $\sum_{m,k} O_{o,mn} O_{d,mp} O_{d,kp} O_{o,kj}$. The arguments given previously for the efficient calculation of the important subset of eigenvectors of A′ also apply here, when the vector **U** is used as the starting vector for the LA. One is thus left with the job of diagonalizing **W**. The application of the LA in this case is less well defined given that there is no well-defined time-independent starting vector to use. Given a congruence between the basis sets, it would seem physically reasonable to use the vector $\mathbf{U}_d$ that is congruent to **U** as the starting vector in the diagonal subspace. Again, the important subset of eigenvectors of **W** could be constructed by the Cullum–Willoughby procedure. Past experience [34, 35, 146] has shown that only a few of the slowest decaying eigenvectors of **W** are significant. Given that the LA is efficient in generating the eigenvalues of smallest real part (cf. Section V.A), it should not require many Lanczos steps to obtain good approximations to them.

Another possible approach is suggested by rewriting Eqs. 194 and 195 in the form

$$S(\tau_1+T+\tau_2) \propto \mathrm{Re}\{\mathbf{U}^{tr}\exp[-\mathbf{A}'\tau_2]\mathbf{V}_{\pi/2}\exp[-\mathbf{W}T]\mathbf{V}_{\pi/2}\exp[-\mathbf{A}'\tau_1]\mathbf{U}\}, \tag{196}$$

where $\mathbf{V}_{\pi/2}$ expresses the effect of the second and third $\pi/2$ pulses. A triple Fourier transform gives

$$S(\omega_1,\omega_2,\omega_T) \propto \mathrm{Re}\{(\mathbf{U}[\omega_2\mathbf{I}+\mathbf{A}']^{-1})\mathbf{V}_{\pi/2}[\omega_T\mathbf{I}+\mathbf{W}]^{-1} \times \mathbf{V}_{\pi/2}([\omega_1\mathbf{I}+\mathbf{A}']^{-1}\mathbf{U})\}. \tag{197}$$

The two expressions in parentheses could, in principle, be calculated using th CG algorithm to solve

$$[\omega_1 \mathbf{I} + \mathbf{A}']\mathbf{z}_1(\omega_i) = \mathbf{U} \tag{198}$$

for $\mathbf{z}_1(\omega_i)$ for a range of values of $\omega_1$ and $\omega_2$ and to store these vectors. Wha remains to be calculated is then

$$S(\omega_1, \omega_2, \omega_T) \propto \mathrm{Re}\{\mathbf{z}^{\mathrm{tr}}(\omega_2)\mathbf{V}_{\pi/2}[\omega_T \mathbf{I} + \mathbf{W}]^{-1}\mathbf{V}_{\pi/2}\mathbf{z}(\omega_1)\}. \tag{199}$$

Now, when a simple congruence exists, the effects of $\mathbf{V}_{\pi/2}$ are simple and lead tc [146] $\chi(\omega_1) = i\sqrt{2}\mathbf{z}^*(\omega_1)$ from the second pulse, and $\mathbf{z}(\omega_2) = i\sqrt{2}\chi(\omega_2)$ from the third pulse so that Eq. 199 becomes

$$S(\omega_1, \omega_2, \omega_T) \propto \mathrm{Re}\{\chi^{\mathrm{tr}}(\omega_2)[\omega_T \mathbf{I} + \mathbf{W}]^{-1}\chi(\omega_1)\}. \tag{200}$$

It would probably still be best to diagonalize $\mathbf{W}$ for the slowest decaying modes by the LA, as discussed in the preceding.

Equations 199 and 200 suggest that $S(\omega_1, \omega_2, \omega_T)$ can be viewed as a matrix of correlation functions of $\chi(\omega_1)$ with $\chi(\omega_2)$ due to the effects of the relaxation

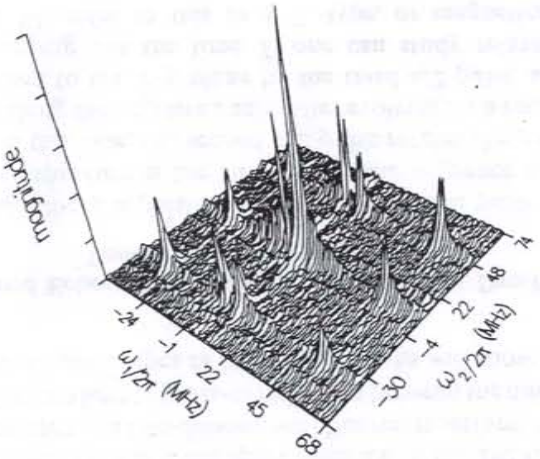



**Figure 17.** Two-dimensional exchange spectrum of 1.17 m*M* PD-TEMPONE spin probe in perdeuterated toluene at 21 °C obtained by Fourier transform methods. Cross-peaks are due to Heisenberg spin exchange inducing magnetization transfer between the primary three lines along diagonal. (Spurious peaks, which do not come at the magnetization transfer positions, are due to residual effects of axial peaks). (From Ref. 34.)



processes associated with **W**. Such a correlation matrix (still a function of $\omega_T$) would contain all of the relevant information obtainable by the stimulated echo sequence. A form intermediate between Eqs. 196 and 197, which we write as $S(\omega_1, \omega_2, T)$, can be used to interpret actual two-dimensional exchange experiments that have been performed both by NMR [155] and by ESR [34, 35] (cf. Figure 17). Work is in progress in making the full three-dimensional experiment a practical reality.

We illustrate some experimental and calculated results in Figure 18 for the stimulated echo and inversion recovery sequences, both of which are magnetization transfer experiments. The inversion recovery experiment is based on the following pulse sequence: The first $\pi$ pulse inverts the $z$

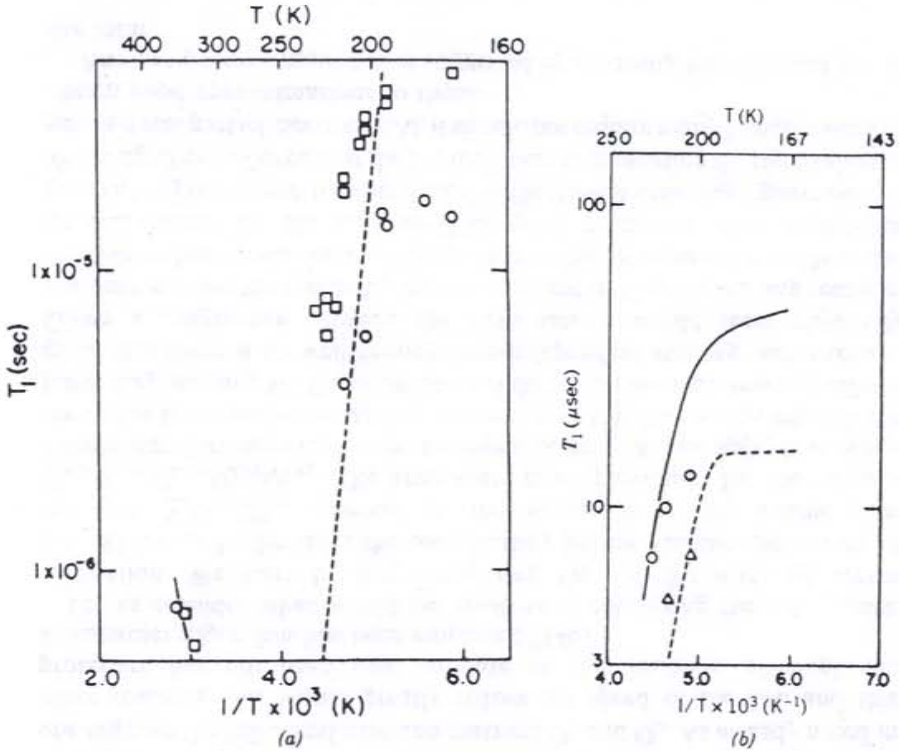



**Figure 18.** Experimental results and calculations on inversion recovery (IR) and stimulated echo (SE) sequences for small spin probe in viscous solvent. (*a*) Apparent $T_1$ versus inverse temperature for IR (squares) and SE (circles) sequences with partial irradiation. Solid line: Measured $T_2$ for center line. Dashed line: Extrapolated fast-motional $\tau_R$ for reference. (*b*) Comparison of experimental and calculated apparent $T_1$. Solid line: Experimental IR. Dashed line: Experimental SE. Circles: Calculated IR. Triangles: Calculated SE. (From Ref. 33.) Courtesy of North-Holland Publ. Co.

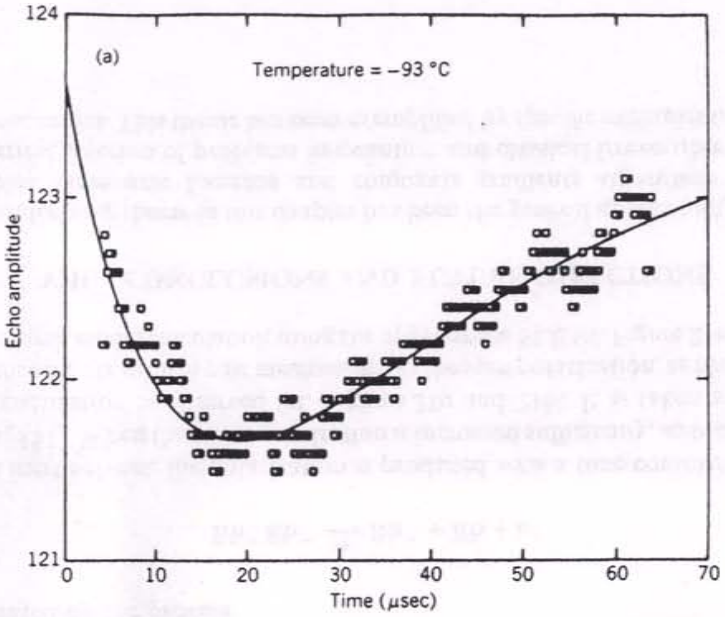



**Figure 19a.** Experimental stepped-field ELDOR echo amplitude versus time between first $\pi$ pulse and $\pi/2$ pulse for small spin probe in viscous solvent at −93 °C. Solid line: Best two exponential fit to data. (From Ref. 29.) Courtesy of North-Holland Publ. Co.

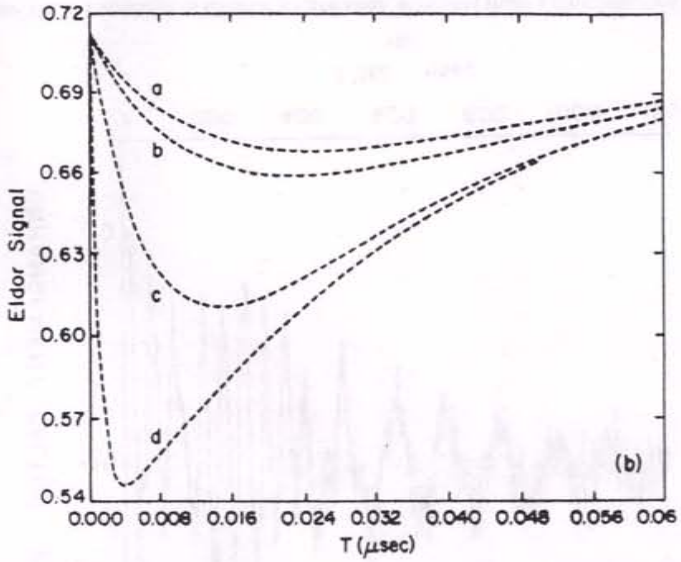



**Figure 19b.** Calculated ELDOR curves for slowly tumbling nitroxide. In sequence a–d orientation-independent nuclear spin flip rate increased, thereby increasing ELDOR effect. (From Ref. 146.)

magnetization; after a time $T$ its magnitude is probed by forming an echo by tipping the magnetization into the $x$–$y$ plane with a $\pi/2$ pulse and then refocusing with another $\pi$ pulse. Closely related to inversion recovery is the echo-ELDOR experiment [29, 156]. It utilizes the same pulse sequence, but the magnetic field is stepped after the first $\pi$ pulse, so the transfer of inverted magnetization to a different spectral position can be studied as illustrated in Figure 19*a* and 19*b*.

The possibilities of a two-dimensional magnetization transfer experiment based on the stimulated echo sequence in studies of motional dynamics are illustrated in Figure 20 for $NO_2$ adsorbed on a VYCOR surface. The anisotropy of the motion can be immediately discerned from the two-dimensional contour plot.

Another class of experiments probe the $z$ magnetization to study the phenomenon of chemically induced dynamic spin polarization [69]. In this experiment, free radicals are produced with a laser pulse, and the development of the spin polarization with its subsequent decay is monitored by a standard

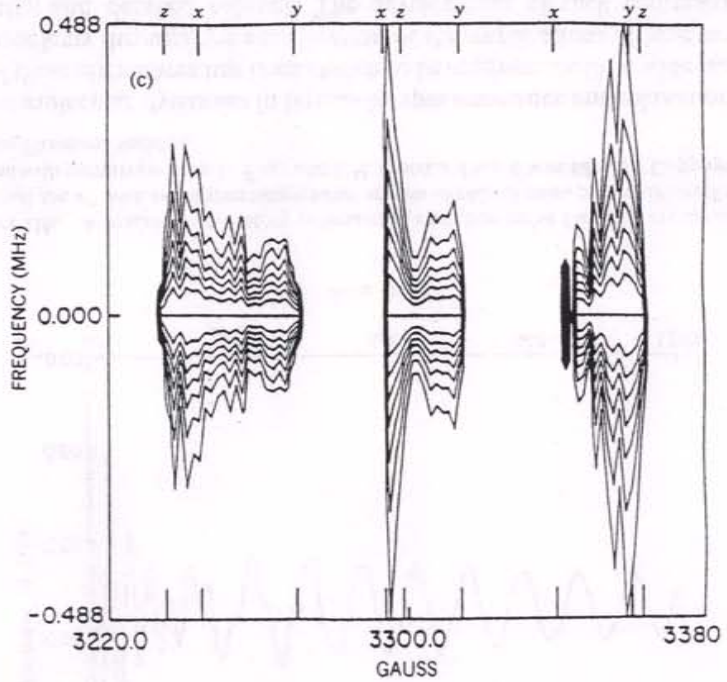



**Figure 20.** Two-dimensional ESE contours from stimulated echo sequence for $NO_2$ on vycor at 35 K showing rates of magnetization transfer. It shows relatively rapid rotation about the molecular $y$ axis (i.e., axis parallel to oxygen–oxygen internuclear vector). (From Ref. 33.) Courtesy of North-Holland Publi. Co

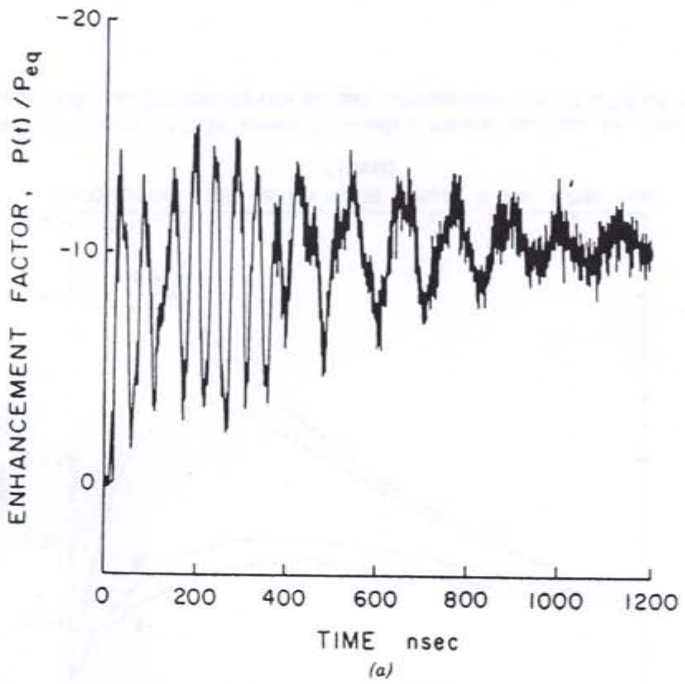



**Figure 21a.** Evolution of CIDEP polarization with time from a photoelectron ejected fro Rb⁻ in THF solvent (From Ref. 157. Reprinted with permission from U. Eliav and J. H. Freed, *Phys. Chem.* **88**, 1277. Copyright 1984 American Chemical Society).

two-pulse spin echo sequence. In the case of a transient photoelectro generated by the process

$$Rb^+Rb^- \xrightarrow{h\nu} Rb^+ + Rb + e^-$$

in an inert solvent, the polarization is produced with a time constant of 10 50 ns [157]. When the Rb concentration is increased sufficiently, an oscillator $z$ magnetization is observed (cf. Figures 21*a* and 21*b*). It is taken as direc evidence for the radical pair mechanism for the spin polarization, as illustrate by a simple model calculation using the appropriate SLE (cf. Figure 21*b*) [157

## VIII. CONCLUSIONS AND FUTURE DIRECTIONS

The underlying theme in this chapter has been the general applicability of th complex symmetric Lanczos and conjugate gradients algorithms to th numerical solution of problems in quantum and classical irreversible statist ical mechanics. This theme has been exemplified by specific examples from th



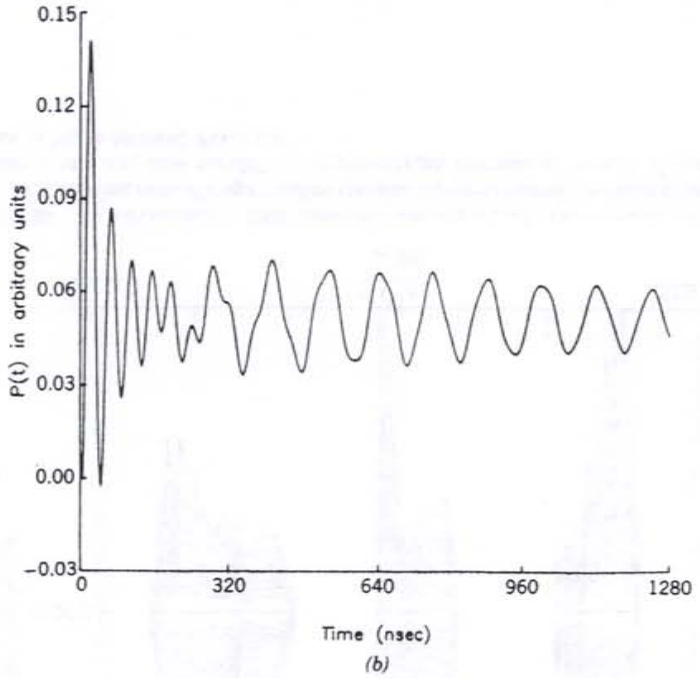



**Figure 21b.** A predicted oscillatory polarization evolution curve for jump reencounters of the Rb· and the $e^-$ with two solvent-dependent species of radical pairs present (From Ref. 157. Reprinted with permission from U. Eliav and J. H. Freed, *J. Phys. Chem.* **88**, 1277. Copyright 1984 American Chemical Society).

study of molecular dynamics in liquids by spin resonance and relaxation. The form of these algorithms has been shown to be appropriate for a wide range of such problems through the examination of the implications of time reversal symmetry and detailed balance. The development of such computational methods, in conjunction with new experimental techniques, should make possible even more powerful methods to study molecular dynamics. Recent work in other areas has further demonstrated the generality and utility of the complex symmetric LA. It has now been used in scattering [158] and optical spectroscopic [159] calculations in conjunction with the complex rotated Hamiltonian method. The properties of the matrices that arise in these studies are closely related to those considered here [126].

We wish to conclude by summarizing some of the ways we believe the application of the LA and CG methods to the class of problems mentioned in the preceding can be improved or enhanced. First, it should be possible to employ a block version [80, 89] of the complex symmetric LA that would be capable of simultaneously calculating the spectral functions corresponding to several auto- and cross-correlation functions. An interesting application of

this algorithm would be in cases where the finite-element discretization is used. A block LA might allow one to very efficiently calculate the cross-correlation of dynamical variables in different localized regions of space. In addition, the block LA and CG methods are more easily adapted to take advantage of the architecture of modern supercomputers. The use of nonorthogonal finite-element functions requires a form of the LA designed to treat the generalized eigenvalue problem. Again, this type of algorithm is well known for real matrices. The nonorthogonal LA has only been used in the study of simple problems to date [44, 160, 161], but applications to more complicated magnetic resonance problems might be fruitful.

The CG method was found to be useful in determining the proper minimum basis set for an accurate numerical calculation of spectra. Future work, possibly relying on artificial intelligence techniques, should be directed toward extending the existing database and developing effective methods of accessing this information. Also, further work on preconditioning and sequence convergence acceleration techniques for the direct calculation of spectra using the CG algorithm would be useful.

Given that the underlying mathematical theory for the complex symmetric LA and CG methods involves the theory of bilinearly metric spaces, further study of the properties of these spaces, especially complex orthogonal spaces, could provide further insight into the behavior of these algorithms and into irreversible statistical mechanics in general.

The analysis of the enormous amount of data generated by modern time domain magnetic resonance techniques has become a complex and sophisticated matter. One method that has recently been proven to be very powerful in time domain ESR experiments is based on using linear prediction in conjunction with the singular-value decomposition. We believe that a Lanczos-based singular-value decomposition [89] can be designed to take advantage of the special structure of the matrix generated by the linear prediction modeling of the signal. The development of such an algorithm has the potential of dramatically reducing the computer time and storage requirements for the routine analysis of experimental data using the linear prediction–singular value decomposition method compared to other algorithms that cannot exploit the special structure of the matrix [215].

An important way in which the comparison of theoretical model calculations and experimental data can be improved is by the automation of the process using nonlinear least-squares fitting procedures. Since this process would involve the calculation of large numbers of approximate spectra, a very efficient LA or CG algorithm that can take advantage of a minimum basis set database is required. This is a very complicated effort because the minimum basis set and optimal number of LA or CG steps might vary widely as the nonlinear fitting algorithm searches for the best set of parameters. This is a problem on which members of our group have recently been making some progress utilizing the CG tridiagonalization procedure.



## APPENDICES

### A. Lanczos Algorithm: A Simple Derivation

A short description of the LA appropriate for matrices of different types is given in this appendix as a reference for readers who are unfamiliar with the algorithm. The derivation given here is slanted toward the generalizations needed for complex symmetric matrices and to simplify the discussions on the connections between the LA method of calculating spectral functions and related methods involving Padé approximants and continued fractions. A convenient starting point in the derivation of the LA is the introduction of a sequence of Krylov vectors generated by successive application of A on $\mathbf{v}$ (see Section V.A),

$$\mathbf{k}_m = \mathbf{A}^{m-1}\mathbf{v} = \mathbf{A}\mathbf{k}_{m-1}, \qquad m = 1, 2, \ldots . \tag{201}$$

The Lanczos basis is generated by the Gram–Schmidt orthonormalization [80, 87] of the sequence of Krylov vectors in the order of their occurrence.

To start the process, assume that the first member in the orthonormal set, $\mathbf{q}_1$, is parallel to the starting vector $\beta_0\mathbf{q}_1 = \mathbf{k}_1 = \mathbf{v}$. The requirement that the new set of vectors be normalized ($\mathbf{q}_i^{tr}\mathbf{q}_i = 1$) implies $\beta_0 = \|\mathbf{v}\|$.

The second member is given by

$$\beta_1\mathbf{q}_2 = \mathbf{k}_2 - \alpha_1\mathbf{k}_1 = (\mathbf{A} - \alpha_1\mathbf{I})\mathbf{q}_1 . \tag{202}$$

The coefficient $\alpha_1$ is determined by multiplying Eq. 202 on the left by $\mathbf{q}_1^{tr}$, whereas $\beta_1$ is chosen to normalize $\mathbf{q}_2$, giving

$$\alpha_1 = \mathbf{q}_1^{tr}\mathbf{A}\mathbf{q}_1 , \tag{203}$$

$$\beta_1 = \|(\mathbf{A} - \alpha_1\mathbf{I})\mathbf{q}_1\| . \tag{204}$$

In a similar fashion, the third vector is

$$\beta_2\mathbf{q}_3 = (\mathbf{A} - \alpha_2\mathbf{I})\mathbf{q}_2 - \gamma_1\mathbf{q}_1 . \tag{205}$$

Again, $\gamma_1$ and $\alpha_2$ are determined from the orthogonality requirements

$$\gamma_1 = \mathbf{q}_1^{tr}\mathbf{A}\mathbf{q}_2 = \beta_1 , \tag{206}$$

$$\alpha_2 = \mathbf{q}_2^{tr}\mathbf{A}\mathbf{q}_2. \tag{207}$$

Inserting these coefficients into Eq. 205 gives a specific case of the general three-term recursion relation that is the heart of the LA,

$$\beta_2\mathbf{q}_3 = (\mathbf{A} - \alpha_2\mathbf{I})\mathbf{q}_2 - \beta_1\mathbf{q}_1. \tag{208}$$

Again, $\beta_2$ is chosen such that $\mathbf{q}_3$ is normalized.

The calculation of the fourth vector demonstrates the appearance of the desired recursion relation. In general, $\mathbf{q}_4$ can be expressed as

$$\beta_3\mathbf{q}_4 = (\mathbf{A} - \alpha_3\mathbf{I})\mathbf{q}_3 - \gamma_2\mathbf{q}_2 - \delta_1\mathbf{q}_1. \tag{209}$$

If Eq. 208 is indeed a prototype of a three-term recursion relation, $\delta_1 = 0$ and $\gamma_2 = \beta_2$. To verify $\delta_1 = 0$, premultiply Eq. 209 by $\mathbf{q}_1^{tr}$ and use the orthogonality property to see that

$$\delta_1 = \mathbf{q}_1^{tr}\mathbf{A}\mathbf{q}_3. \tag{210}$$

Using Eq. 202, this can be rewritten as

$$\delta_1 = (\beta_1\mathbf{q}_2 + \alpha_1\mathbf{q}_1)^{tr}\mathbf{q}_3, \tag{211}$$

which vanishes by the orthogonality of the $\mathbf{q}$'s. In addition, premultiplying Eqs. 208 and 209 by $\mathbf{q}_3^{tr}$ gives $\beta_2 = \gamma_2$.

The same behavior is observed for all further vectors. Therefore, the three-term recursion relation

$$\beta_m\mathbf{q}_{m+1} = (\mathbf{A} - \alpha_m\mathbf{I})\mathbf{q}_m - \beta_{m-1}\mathbf{q}_{m-1} \tag{212}$$

can be used to generate all successive $\mathbf{q}$'s so long as the $\beta$'s are nonzero. The orthonormal basis of $\mathcal{K}_m(\mathbf{A}, \mathbf{v})$ generated in this manner is called the Lanczos basis, and the individual vectors are called Lanczos vectors.

It is informative to rewrite these relations in a matrix form. The transformation matrix $\mathbf{Q}_m$, whose $j$th column is given by the elements of $\mathbf{q}_j$ for $1 \leqslant j \leqslant m$, transforms the original $N \times N$ matrix $\mathbf{A}$ into an $m \times m$ real symmetric tridiagonal matrix $\mathbf{T}_m$,

$$\mathbf{T}_m = \mathbf{Q}_m^{tr}\mathbf{A}\mathbf{Q}_m. \tag{213}$$

The nonzero matrix elements of $\mathbf{T}_m$ on and above the diagonal are given by

$$(T_m)_{i,i} = \alpha_i, \tag{214}$$



$$(T_m)_{i,i+1} = \beta_i. \tag{215}$$

In addition, the orthonormalization process can be rewritten as a matrix equation relating $\mathbf{Q}_m$ to the matrix of Krylov vectors $\mathbf{K}_m$ and the moments (see Appendix B)

$$\mu_j = \mathbf{v}^{tr}\mathbf{A}^{j-1}\mathbf{v} \tag{216}$$

using the relations

$$\begin{aligned}
\beta_0\mathbf{q}_1 &= \mathbf{k}_1, \\
\beta_1\mathbf{q}_2 &= \mathbf{k}_2 - (\mathbf{k}_2^{tr}\mathbf{k}_1)\mathbf{k}_1 \\
&= \mathbf{k}_2 - \mu_1\mathbf{k}_1 \\
\beta_2\mathbf{q}_3 &= \mathbf{k}_3 - (\mathbf{k}_3^{tr}\mathbf{k}_2)\mathbf{k}_2 - (\mathbf{k}_3^{tr}\mathbf{k}_1)\mathbf{k}_1 \\
&= \mathbf{k}_3 - \mu_3\mathbf{k}_2 - \mu_2\mathbf{k}_1,
\end{aligned}$$

and so forth. These relations follow directly from the definitions of the quantities involved. In matrix form these equations read

$$\mathbf{Q}_m\mathbf{B}_m = \mathbf{K}_m\mathbf{U}_m, \tag{217}$$

where $\mathbf{B} = \mathrm{diag}(\beta_0, \beta_1, \beta_2, \ldots)$ and the matrix $\mathbf{U}_m$ is the upper triangular matrix

$$\mathbf{U}_m = \begin{bmatrix}
1 & -\mu_1 & -\mu_2 & -\mu_3 & -\mu_4 \cdots \\
0 & 1 & -\mu_3 & -\mu_4 & -\mu_5 \cdots \\
0 & 0 & 1 & -\mu_5 & -\mu_6 \cdots \\
0 & 0 & 0 & 1 & -\mu_7 \cdots \\
0 & 0 & 0 & 0 & 1 \\
\vdots & & \vdots & \vdots & \ddots
\end{bmatrix}. \tag{218}$$

Since $\mathbf{U}_m$ is nonsingular (all its eigenvalues are equal to unity), Eq. 217 can be reexpressed as

$$\mathbf{K}_m = \mathbf{Q}_m(\mathbf{B}_m\mathbf{U}_m^{-1}) = \mathbf{Q}_m\mathbf{R}_m. \tag{219}$$

In this equation $\mathbf{U}_m^{-1}$ and $\mathbf{R}_m$ are upper triangular since $\mathbf{U}_m$ has this property. Therefore, Eq. 219 represents the factorization of the matrix of Krylov vectors into a product of a real matrix with orthonormal columns $\mathbf{Q}_m$ and an upper triangular matrix $\mathbf{R}_m = \mathbf{B}_m\mathbf{U}_m^{-1}$. Furthermore, if $\mathbf{B}_m$ is nonsingular, so is $\mathbf{R}_m$.

Matrix factorizations such as that given by Eq. 219 are of great importance in modern numerical linear algebra. The existence of factorizations of this type for arbitrary real or complex matrices is guaranteed by the following theorem [87].

***Theorem 1 (QR Factorization)*** *For any given $N \times M$ complex matrix* **A** *($N \geq M$), there exists a matrix* **Q** *of the same dimensions with orthonormal columns ($\mathbf{Q}^\dagger\mathbf{Q} = \mathbf{I}_N$) and an upper triangular $M \times M$ matrix* **R** *such that* $\mathbf{A} = \mathbf{QR}$. *If* **A** *is real,* **Q** *and* **R** *can also be chosen to be real.*

The Gram–Schmidt orthonormalization procedure can always be used to accomplish this factorization, though there are also many other ways [80].

The extension of the LA to handle Hermitian matrices can be derived in a straightforward manner by replacing the transposition operations with Hermitian conjugation in the previous equations and taking advantage of the symmetry of the usual inner product in a unitary space (see Appendix C). The desired decomposition of the matrix of Krylov vectors is ensured by the *QR* factorization theorem above. The result of these operations gives the Lanczos recursion relation for a general Hermitian matrix **A**:

$$\beta_m \mathbf{q}_{m+1} = (\mathbf{A} - \alpha_m \mathbf{I})\mathbf{q}_m - \beta_{m-1}\mathbf{q}_{m-1}, \tag{220}$$

where $\alpha_i = \mathbf{q}_i^\dagger \mathbf{A} \mathbf{q}_i$ and $\beta_i$ is again chosen to normalize $\mathbf{q}_{i+1}$. The reduction of **A** to a tridiagonal matrix is given by

$$\mathbf{T}_m = \mathbf{Q}_m^\dagger \mathbf{A} \mathbf{Q}_m, \tag{221}$$

where $\mathbf{Q}_m$ satisfies $\mathbf{Q}_m^\dagger \mathbf{Q}_m = \mathbf{I}_m$. It is easy to see from the Eq. 220 that an arbitrary Hermitian matrix will still be reduced to a real symmetric tridiagonal form by the general Hermitian LA. It is important to note that the form of the recursion relation is the same as before. The only way the complex nature of the matrix enters is in the definition of the scalar product.

The extension of the LA to handle non-Hermitian, complex symmetric matrices is not so simple. One way to summarize these difficulties is to say that although the *QR* factorization of the matrix of Krylov vectors generated is natural in the Hermitian case, it is fundamentally inappropriate for the Krylov matrices generated by a complex symmetric matrix. This fact is emphasized by examining Eq. 221 for a case where the Hermitian matrix **A** has no degenerate eigenvalues such that the matrix can be completely reduced to a tridiagonal form [89, 80] of the same dimension by the LA. In this instance, the matrix **Q** becomes unitary and Eq. 221 is a statement that the Hermitian matrix **A** is unitarily similar to Hermitian matrix **T**. However, if two complex symmetric



matrices are similar, they must be similar via a complex orthogonal transformation [86, 87]. Thus, a generalization of the LA along different lines must be sought. It turns out that the required generalization amounts to a redefinition of the inner product.

To develop a complex symmetric LA, a different factorization of the matrix of Krylov vectors must be adopted. This amounts to developing an analog of the Gram–Schmidt orthonormalization procedure, which will naturally lead to complex orthogonal matrices rather than unitary matrices. Choudhury and Horn have studied this problem in detail and have shown when such a factorization is possible [90]. In contrast to the Hermitian case, it turns out that there are instances where this generalized Gram–Schmidt procedure fails even where a related factorization of the Krylov matrix exists. In the present context it is the generalization of the Gram–Schmidt procedure that is the more important. The statement of the relevant definitions and theorem and further discussion can be found in Appendix C. Using the result of Theorem 2 in Appendix C, the Gram–Schmidt-like orthogonalization of the Krylov vectors generated by a complex symmetric matrix **A** and an arbitrary normalized starting vector depends on the determinants

$$\det(\mathbf{K}_m^{\mathrm{tr}}\mathbf{K}_m) \tag{222}$$

being different from zero for all $m$ in question. One of these determinants being zero implies that the associated Lanczos vector is nonzero but has zero norm, and the matrix $\mathbf{Q}_m$ of Lanczos vectors cannot satisfy $\mathbf{Q}_m^{\mathrm{tr}}\mathbf{Q}_m = \mathbf{I}_m$. If none of these determinants are nonzero, the complex symmetric LA behaves in the same fashion as the Hermitian LA [80, 89].

Since the columns of $\mathbf{K}_m$ are the Krylov vectors, the product $\mathbf{K}_m^{\mathrm{tr}}\mathbf{K}_m$ can be rewritten in terms of the moments. Using the definition of the Krylov vectors (see Eq. 201), it is easy to show that

$$\Delta_m = \det(\mathbf{K}_m^{\mathrm{tr}}\mathbf{K}_m) = \det \begin{bmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_m \\ \mu_1 & \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{m+1} \\ \mu_2 & \mu_3 & \mu_4 & \mu_5 & \cdots & \mu_{m+2} \\ \mu_3 & \mu_4 & \mu_5 & \mu_6 & \cdots & \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \mu_m & \mu_{m+1} & \mu_{m+2} & \cdots & \cdots & \mu_{2m} \end{bmatrix}. \tag{223}$$

In Appendix B determinants of this type will also arise in discussing continued fraction and Padé approximants to the spectrum.

The requirement that the determinants in Eq. 223 be nonzero seems to be

the only constructive criterion for the existence of a well-behaved complex symmetric LA in exact arithmetic. A derivation of a similar criterion for Hermitian matrices is given by Householder [162]. The result given here is actually a special case of Householder's analysis of the biorthogonal LA. Other criteria can be found for the existence of a well-behaved complex symmetric LA, such as the requirement that the minimal polynomial of the matrix with respect to the starting vector be a product of distinct linear factors [1]. This type of criterion is not constructive in the sense that if one already knew the minimal polynomial, the problem would already be solved!

### B. Continued Fractions, Padé Approximants, and the Lanczos Algorithm

The essence of the Lanczos method of calculating spectral functions is to use the sequence of symmetric tridiagonal matrices generated by the LA to define a sequence of rational, or continued-fraction, approximants to the desired spectral function. The purpose of this appendix is to develop the theory of these approximants and to demonstrate the close connection of this method to related methods involving the calculation of continued fractions and Padé approximants directly from the moment expansion. The focus of this discussion will be on cases where the time evolution of the system is governed by a non-Hermitian complex symmetric matrix $\mathbf{A}_N$, since results similar to those developed here are well known for the Hermitian case.

The generalized spectral density

$$J^{(N)}(z) = \mathbf{u}^{\dagger}[z\mathbf{I}_N + \mathbf{A}_N]^{-1}\mathbf{v} \tag{224}$$

will serve as a reference point in this discussion. It is assumed that the projections $P_N$ defining the truncated matrix $\mathbf{A}_N = P_N\mathbf{A}P_N$ have been chosen such that $J^{(N)}(z)$ is a very good approximation to the "true" spectral function and that the components of $\mathbf{u}$ are the complex conjugates to the components of $\mathbf{v}$ in the orthonormal basis set chosen in which $\mathbf{A}_N$ is complex and symmetric (see Sections VI.B.1 and VI.D) [3]. With these assumptions, $J^{(N)}(z)$ can be expressed as

$$J^{(N)}(z) = \mathbf{v}^{tr}[z\mathbf{I}_N + \mathbf{A}_N]^{-1}\mathbf{v}. \tag{225}$$

#### 1. *Continued Fraction Approximants Derived from the Lanczos Algorithm*

The rational approximants $J_n^{(N)}(z)$ to $J^{(N)}(z)$ are defined as

$$J_n^{(N)}(z) = (\mathbf{v}^{tr}\mathbf{Q}_n)\mathbf{Q}_n^{tr}[z\mathbf{I}_N + \mathbf{A}_N]^{-1}\mathbf{Q}_n(\mathbf{Q}_n^{tr}\mathbf{v}) \tag{226}$$

$$= (\mathbf{Q}_n^{tr}\mathbf{v})^{tr}[z\mathbf{I}_n + \mathbf{T}_n]^{-1}(\mathbf{Q}_n^{tr}\mathbf{v}). \tag{227}$$



Since the matrix $\mathbf{Q}_n$ satisfies $\mathbf{Q}_n^{tr}\mathbf{Q}_n = \mathbf{I}_n$ and the elements in the first column of $\mathbf{Q}_n$ are just the components of $\mathbf{v}$, the $n$th approximant to $J^{(N)}(z)$ is just the (1, 1) matrix element of the inverse of $[z\mathbf{I}_n + \mathbf{T}_n]$ (see Sections II and II.A):

$$J_n^{(N)}(z) = [z\mathbf{I}_n + \mathbf{T}_n]^{-1}_{1,1}. \tag{228}$$

A similar result can be derived for the Hermitian case if the complex orthogonal matrix $\mathbf{Q}_n$ is replaced with the appropriate unitary matrix

To systematically develop these ideas in a manner that will simplify the discussion of the relationship between the rational approximants, continued fractions, and Padé approximants to the spectral function it is useful to introduce the determinants $D_{l,m}^n$ of the diagonal blocks of $[z\mathbf{I}_n + \mathbf{T}_n]$,

$$D_{l,m}^n(z) = \det\begin{bmatrix} z+\alpha_l & \beta_l & & & \\ \beta_l & z+\alpha_{l+1} & \beta_{l+1} & & \\ & \beta_{l+1} & z+\alpha_{l+2} & \ddots & \\ & & \ddots & \ddots & \beta_{m-1} \\ & & & \beta_{m-1} & z+\alpha_m \end{bmatrix}, \tag{229}$$

where it is assumed that $m \geqslant l$.

In terms of these determinants, the application of Cramer's rule to Eq. 228 gives

$$J_n^{(N)}(z) = \frac{D_{2,n}^n(z)}{D_{1,n}^n(z)} \quad \text{for} \quad n \geqslant 2. \tag{230}$$

It is clear from Eq. 230 that $J_n^{(N)}(z)$ is indeed a rational function since $D_{1,n}^n(z)$ and $D_{2,n}^n$ are polynomials of order at most $n$ and $n-1$, respectively, in the complex variable $z$. A recurrence relation for the determinants $D_{1,n+1}^{n+1}(z)$ and $D_{2,n+1}^{n+1}(z)$ can easily be derived by expanding the determinants in question using Laplace's method along the last row or column.

The sequence of rational approximants can be converted to a continued fraction by expanding the determinants in a slightly different fashion. The expansion of $D_{1,n}^n(z)$ about its first row or column gives the result

$$D_{1,n}^n(z) = (z+\alpha_1)D_{2,n}^n(z) - \beta_1^2 D_{3,n}^n(z). \tag{231}$$

Inserting this expansion into the denominator of Eq. 230 and dividing both numerator and denominator by $D_{2,n}^n(z)$ gives

$$J_n^{(N)}(z) = \frac{D_{2,n}^n(z)}{(z+\alpha_1)D_{2,n}^n - \beta_1^2 D_{3,n}^n(z)} \tag{232}$$

$$= \left( z + \alpha_1 - \beta_1^2 \frac{D_{3,n}^n(z)}{D_{2,n}^n(z)} \right)^{-1}. \tag{233}$$

This process can be repeated for the ratio of determinants occurring in the denominator to yield

$$J_n^{(N)}(z) = \left( z + \alpha_1 - \frac{\beta_1^2}{z + \alpha_2 - \beta_2^2 \dfrac{D_{4,n}^n(z)}{D_{3,n}^n(z)}} \right)^{-1}. \tag{234}$$

The recursive nature of the expansion of the rational approximant into a continued fraction is now obvious. Using the standard notation for continued fractions, the complete expansion of $J_n^{(N)}(z)$ can be written as

$$J_n^{(N)}(z) = \frac{1}{z + \alpha_1} - \frac{\beta_1^2}{z + \alpha_2} - \frac{\beta_2^2}{z + \alpha_3} - \cdots \frac{\beta_{n-2}^2}{z + \alpha_{n-1}} - \frac{\beta_{n-1}^2}{z + \alpha_n}. \tag{235}$$

A few definitions are in order at this time to establish the terminology and notation for continued fractions used here.

***Definition 1 (Elements of Continued Fraction)*** *The complex numbers $a_m$ and $b_m$ are the called the elements of the (infinite) continued fraction*

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} + \frac{a_3}{b_3} + \cdots. \tag{236}$$

***Definition 2 (Partial Numerators and Denominators)*** *The complex numbers $a_m$ and $b_m$ are the $m^{th}$ partial numerator and partial denominator, respectively, of the continued fraction Eq. 236.*

***Definition 3 (Convergents)*** *The continued fraction*

$$C_m = \frac{a_1}{b_1} + \frac{a_2}{b_2} + \frac{a_3}{b_3} + \cdots \frac{a_m}{b_m} \tag{237}$$

*is called the $m^{th}$ convergent of the continued fraction Eq. 236. The $m^{th}$ convergent can be collapsed into a single fraction by simple arithmetic. The numerator $A_m$ and denominator $B_m$ of this fraction are simply called the $m^{th}$ numerator and denominator of Eq. 236. In terms of these quantities, the $m^{th}$ convergent is just*

$$C_m = \frac{A_m}{B_m}. \tag{238}$$



The approximants $J_n^{(N)}(z)$ are the convergents of the continued fraction Eq. 235.

Some of the most basic and useful results of the theory of continued fractions are the fundamental three-term recurrence relations for the numerators and denominators:

$$A_{m+1} = b_{m+1} A_m + a_{m+1} A_{m-1}, \tag{239}$$

$$B_{m+1} = b_{m+1} B_m + a_{m+1} B_{m-1}, \tag{240}$$

for $m = 0, 1, 2, \ldots$. The value of any convergent is easily calculated using these recurrence relations with the initial conditions

$$A_{-1} = 1, \qquad A_0 = 0, \tag{241}$$

$$B_{-1} = 0, \qquad B_0 = 1. \tag{242}$$

These recurrence relations are numerically unstable in most cases but are quite useful for analytical studies.

***Definition 4 (Convergence of Continued Fraction)*** *A continued fraction such as Eq. 236 is said to converge if the limit of its sequence of convergents*

$$\lim_{m \to \infty} \frac{A_m}{B_m} = \lim_{m \to \infty} C_m \tag{243}$$

*converges to a finite number C and only a finite number of the denominators vanish. The number C is called the value of the continued fraction. A continued fraction that does not satisfy these requirements is said to diverge.*

The other basic results from the general theory of continued fractions needed here are those concerned with equivalent continued fractions and equivalence transformations.

***Definition 5 (Equivalent Continued Fractions)*** *The continued fraction with elements $a_m$ and $b_m$ and convergents $C_m$ is said to be equivalent to another continued fraction with elements $a'_m$ and $b'_m$ and convergents $C'_m$ if $C_m = C'_m$ for all m.*

***Definition 6 (Equivalence Transformations)*** *An equivalence transformation is a mapping of the elements of one continued fraction into an equivalent continued fraction defined by a sequence of nonzero complex constants $r_m$:*

$$a'_m = r_m r_{m-1} a_m, \qquad m = 0, 1, 2, \ldots, \tag{244}$$

$$b'_m = r_m b_m, \qquad m = 1, 2, 3, \ldots, \tag{245}$$

*where* $r_0 = 1$.

It is not hard to show that two continued fractions are equivalent if and only if they are related by an equivalence transformation [163].

Continued fractions of the form given in Eq. 235 have been extensively studied by Wall and co-workers and are known as $J$-fractions [163–167]. The advantage of identifying the continued fractions derived from the tridiagonal matrices generated by the LA as $J$-fractions is that the general theory of $J$-fractions can then be used to treat approximants to the spectral functions in both the Hermitian and complex symmetric cases in a uniform fashion.

Of special interest in physical applications is the class of positive definite $J$-fractions. The class of positive definite $J$-fractions is characterized by the requirement that the real-valued quadratic form derived from the imaginary parts of the elements of the continued fraction is nonnegative for all $n$ [164].

***Definition 7 (Positive Definite J-Fractions)*** *A continued fraction of the form*

$$\frac{1}{b_1 + \zeta} - \frac{a_1^2}{b_2 + \zeta} - \frac{a_2^2}{b_3 + \zeta} - \cdots \tag{246}$$

*is called positive definite J-fraction if, for all M and all real values of* $y_1$, $y_2$, $y_3, \ldots$,

$$\sum_{k=1}^{M} \delta_k y_k^2 - 2 \sum_{k=1}^{M-1} \gamma_k y_k y_{k+1} \geqslant 0, \tag{247}$$

*where* $\delta_k = \mathrm{Im}\,(b_k)$ *and* $\gamma_m = \mathrm{Im}\,(a_k)$.

To simplify the connections to the mathematical literature, it is useful to perform an equivalence transformation on the continued-fraction expansion of the spectral function derived from the Lanczos tridiagonal matrices of the form given in Eq. 235 to yield the alternative form



$$J_n^{(N)}(z) = -iJ_n^{(N)}(\zeta) = \frac{1}{b_1 + \zeta} - \frac{a_1^2}{b_2 + \zeta} - \frac{a_2^2}{b_3 + \zeta} - \frac{a_3^2}{b_4 + \zeta} - \cdots \frac{a_n^2}{b_n + \zeta}, \tag{248}$$

where $b_m = i\alpha_m$, $a_m = i\beta_m$, and $\zeta = -iz$. The continued fraction on the right side of Eq. 248 is a $J$-fraction as defined by Wall [164].

The numerator and denominator of a positive definite $J$-fraction can be found by inserting the elements of the continued fraction Eq. 246 into Eqs. 239 and 240 and the initial conditions in Eqs. 241 and 242:

$$A_n(\zeta) = (b_n + \zeta) A_{n-1} - a_{n-1}^2 A_{n-2}(\zeta), \tag{249}$$

$$B_n(\zeta) = (b_n + \zeta) B_{n-1} - a_{n-1}^2 B_{n-2}(\zeta), \tag{250}$$

where $a_0$ is defined to be unity.

To study the convergence and analytical behavior of infinite $J$-fractions, it is convenient to introduce the two series

$$X_{n+1}(\zeta) = \frac{A_n(\zeta)}{\prod_{j=1}^{n} a_j},$$

$$Y_{n+1}(\zeta) = \frac{B_n(\zeta)}{\prod_{j=1}^{n} a_j},$$

and the infinite series

$$\sum_{j=1}^{\infty} |X_p(0)|^2, \tag{251}$$

$$\sum_{j=1}^{\infty} |Y_p(0)|^2. \tag{252}$$

The polynomials $X_n(\zeta)$ and $Y_n(\zeta)$ each satisfy a three-term recursion relation. If either or both series in Eq. 251 or 252 diverges it is said that the determinate case holds for the continued fraction Eq. 246; otherwise, the indeterminate case holds. It appears that only the indeterminate case is important for the study of spectral functions.

The interest in the class of positive definite $J$-fractions for which the determinate case holds arises from the following observations [163, 164]:

- The denominators of such continued fractions are nonzero for all complex values of $\zeta$ in the open upper half-plane $\mathrm{Im}(\zeta) > 0$.
- They converge uniformly over every closed bounded region in the upper half-plane, and their values are analytic functions of $\zeta$ there.
- Explicit truncation error bounds and inclusion regions for the value of an infinite continued fraction can be derived in terms of the elements of the truncated continued fraction.
- The values of the convergents of a continued fraction of this class satisfy $\mathrm{Im}\{C_m(\zeta)\} < 0$ for all $\zeta$ such that $\mathrm{Im}\{\zeta\} > 0$.

A comparison of these properties with the requirements that must be satisfied for a spectral function to be physically meaningful [51, 108, 109, 168] shows that finite positive definite $J$-fractions possess the mathematical features necessary to properly approximate a spectral function. The physical interpretation of the analytic behavior of $J^{(N)}(z)$ in the upper half-plane is that adding a real constant matrix to $\mathbf{A}_N$ is equivalent to increasing all the relaxation rates by that same amount, which should not alter the basic mathematical structure of the approximation. A related idea has been used by Dammers [49] as a test of the numerical stability of various methods of calculating magnetic resonance spectra (see what follows).

From the point of view of approximating the true spectral function by the sequence of continued-fraction approximants generated by the LA, the only difference is that the Hermitian LA gives rise to continued-fraction approximants with purely real elements, whereas the complex symmetric LA gives rise to complex elements. The previous discussion on positive definite $J$-fraction approximants to spectral functions applies equally well to the complex symmetric and Hermitian LA.

### 2. *Spectral Functions as Moment Problems*

One systematic way to develop an approximation to $J(z)$ is to use a binomial expansion of Eq. 225:

$$J^{(N)}(z) = \frac{1}{z}\mathbf{v}^{\mathrm{tr}}\left[\mathbf{I}_N - \frac{\mathbf{A}_N}{z} + \left(\frac{\mathbf{A}_N}{z}\right)^2 - \left(\frac{\mathbf{A}_N}{z}\right)^3 + \cdots\right]\mathbf{v} \tag{253}$$

$$= \frac{1}{z}\left\{\mathbf{v}^{\mathrm{tr}}\mathbf{v} - \mathbf{v}^{\mathrm{tr}}\frac{\mathbf{A}_N}{z}\mathbf{v} + \mathbf{v}^{\mathrm{tr}}\left(\frac{\mathbf{A}_N}{z}\right)^2\mathbf{v} - \mathbf{v}^{\mathrm{tr}}\left(\frac{\mathbf{A}_N}{z}\right)^3\mathbf{v} + \cdots\right\} \tag{254}$$

$$= \frac{1}{z}\sum_{n=0}^{\infty}\left(\frac{-1}{z}\right)^n \mu_{n+1}, \tag{255}$$



where quantities $\mu_{n+1} = \mathbf{v}^{\mathrm{tr}}\mathbf{A}_N^n\mathbf{v}$ are known as the moments of $\mathbf{A}$ with respect to $\mathbf{v}$, (cf. Eq. 216). The expansion in Eq. 255 is commonly referred to as the moment expansion. Numerical results using this expansion directly are quite disappointing: The convergence is slow, and the truncated expansion does not have the proper analytic behavior as a function of $z$.

The generalized, or relaxation, moment expansion Eq. 255 can be used to calculate a continued-fraction approximant with better convergence properties. Several groups have addressed these issues [48, 169, 170] from the point of view of calculating magnetic resonance spectra by developing formulas for the moments $\mu_n$, for $n = 0, 1, 2, \ldots, N$, and using the moments to calculate the elements of a continued fraction of the type of Eq. 248. This seems to be a useful approach for simple problems where tractable formulas for the moments are available. The formulas needed to calculate the moments given by Giordano et al. [48] are extremely complicated even for the simple cases (e.g., where the $g$ and $A$ tensors of the spin probe are axially symmetric in the principal axis frame of the diffusion tensor and nonsecular terms have been omitted from the spin Hamiltonian). Nevertheless, this is useful for such simple problems, especially if analytical formulas are needed. Very often nonaxially symmetric magnetic tensors are required to quantitatively fit experimental data [12]. Experiments have also shown that much more information is available if the spectra of oriented radicals in liquid crystalline phases are studied as a function of the tilt angle between the director and the static field [20, 21]. The lack of generality, together with the fact that extended precision computer arithmetic is often needed to stably compute the elements of the continued fraction [48, 169], suggest that the Lanczos method is more appropriate than the relaxation moment method for the routine analysis of complicated experimental spectra.

An interesting connection exists between the $J$-fraction approximants of the moment expansion (Eq. 255), the existence of a set of polynomials orthogonal relative to the sequence of moments [164, 171], and the determinants $\Delta_m$ (Eq. 223). First, it is necessary to define what is meant by the orthogonality of a set of complex polynomials relative to a complex sequence.

***Definition 8 (System of Polynomials Orthogonal Relative to a Sequence)*** *A sequence of complex polynomials $Q_n(x)$ in the complex variable $x$,*

$$Q_n(x) = \sum_{m=0}^{n} q_m^{(n)} x^m, \qquad n = 0, 1, 2, \ldots, \tag{256}$$

*is said to be orthogonal relative to a sequence of complex constants*

$\{c_0, c_1, c_2, \ldots\}$ *if*

$$\sum_{i=0}^{m} \sum_{j=0}^{n} q_i^{(m)} q_j^{(n)} c_{i+j} \tag{257}$$

*is nonzero if and only if* $m = n$.

A necessary and sufficient condition for a set of $M+1$ such polynomials to exist that are orthogonal relative to the sequence of moments (Eq. 255) is that the determinants $\Delta_m$ be nonzero for all $m < M-1$ [164]. Moreover, it can be shown that these polynomials are identical to the denominator polynomials given by Eq. 250 of the continued fraction given in Eq. 248. Thus, it is evident that the relaxation moment method is equivalent to the Lanczos method.

An alternative approach involving continued fractions is to seek a classical moment expansion of the absorption spectrum or spectral density. There is a vast mathematical literature on classical moment problems and their relationship to the theory of general orthogonal polynomials, continued fractions, and Padé approximants [163, 164, 172, 173]. This viewpoint differs significantly from the LA or relaxation moment approaches, where the continued-fraction expansion of a generalized spectral function is studied. The difference between these approaches lies in the fact that the absorption spectrum or spectral density can be treated as a *real* nonnegative function over the entire range of frequencies, whereas the related generalized spectral function is inherently complex valued.

In the classical moment method, the quantities of interest are the moments of the frequency with respect to the absorption spectrum or spectral density, $I(\omega)$,

$$\langle \omega^n \rangle = \int_{-\infty}^{\infty} \omega^n I(\omega)\, d\omega, \tag{258}$$

and not the moments of the operator $\mathbf{A}_N$, as in Eq. 255.

Gordon [174] has studied the response of a spin system to a pulsed excitation as a classical Hamburger moment problem [172, 173] and has derived truncation error bounds and inclusion regions similar to those alluded to earlier in connection with positive definite $J$-fractions. This is not surprising given that the solution to the Hamburger moment problem can be expressed as a real, positive definite $J$-fraction [164]. Lado, Memory, and Parker have used similar techniques in conjunction with memory function ideas [175]. Lonke has also derived related results for the spectral density of excitations in a many-body system of fermions [176]. A nice review of the application of the theory of continued fractions to physical problems involving relaxation is given by Grosso and Pastori Parravicini [170].



### 3. *Padé Approximants to Spectral Functions*

Padé approximants [164, 178, 179] have also been used to calculate slow-motional ESR spectra by Dammers, Levine, and. Tjon [46, 49]. These calculations involve the determination of the coefficients of a certain sequence of rational approximants to the moment expansion of the spectral function (see Eq. 255). A Padé approximant is a rational approximant that satisfies the restrictions stated in the following definition.

***Definition 9 (Padé Approximants)*** *Let* $A(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \cdots$ *be a given formal power series and* $P_L(z)$ *and* $Q_M(z)$ *be polynomials of degree at most* $L$ *and* $M$, *respectively, where* $Q_M(0) = 1$. *Let*

$$C(L/M) = det \begin{bmatrix} a_{L-M+1} & a_{L-M+2} & a_{L-M+3} & \cdots & a_L \\ a_{L-M+2} & a_{L-M+3} & a_{L-M+4} & \cdots & a_{L+1} \\ a_{L-M+3} & a_{L-M+4} & a_{L-M+5} & \cdots & a_{L+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_L & a_{L+1} & a_{L+2} & \cdots & a_{L+M-1} \end{bmatrix} \neq 0. \tag{259}$$

*If* $A(z) - P_L(z)/Q_M(z) = \mathcal{O}(z^{L+M+1})$, *then* $[L/M]_A(z) = P_L(z)/Q_M(z)$ *is called the* $L/M$ *Padé approximant to* $A(z)$.

Since the Padé approximants to the moment expansion might exist for arbitrary $L$ and $M$, it is necessary to determine which of these approximants have the proper analytical structure as a function of the complex variable $z$ to represent physically reasonable spectral functions. The conclusions given by Dammers from these analyses, together with the examination of model calculations and the connections of this method to Mori's projection operator method, was that the $[M-1/M]$ Padé approximant to $zJ^{(N)}(z)$ is the proper choice. (See also [3]).

In light of the selection of the class of $[M-1/M]$ Padé approximants as the preferred set, it is interesting to examine the connections between this class of rational approximants and the continued-fraction approximants from the LA or relaxation moment methods. In particular, the determinant $C(M-1/M)$ occurring in the definition of the $[M-1/M]$ Padé approximant (Eq. 259) is identical to the determinant $\Delta_M$ (Eq. 223) occurring in the analyses of the LA and the relaxation moment method. Thus, the existence of the class of $[M-1/M]$ approximants is equivalent to saying that the related LA is well behaved or that the elements of the continued fraction can be calculated from the moments $\mu_i$.

The other way in which Padé approximants have been used in the

calculation of magnetic resonance spectra is to accelerate the convergence of the sequence of estimates of the amplitude of a spectrum given by the CG algorithm at a particular field position [4]. Because of the equivalence of the results of the LA and CG algorithm, it is possible to identify the sequence of estimates given by the CG method with the values of the successive convergents of the continued fraction given by the LA evaluated at the same point. Given the set $\{C_m\}$ of CG estimates, the series

$$C_0 + \sum_m (C_{m+1} - C_m)x^m \tag{260}$$

can be constructed. Inserting the eigenvector–eigenvalue decomposition of the matrix $\mathbf{A}_N$ in Eq. 255 for the moment expansion shows that the moment expansion has a form that suggests that a Padé or generalized Shanks resummation technique might accelerate its convergence [130]. It also follows from the preceding discussion that the sequence of values of the convergents of continued-fraction expansion and the series Eq. 260 also have this property. The literature on Padé or generalized Shanks resummation techniques and related topics such as the epsilon algorithm is far too voluminous and technical to be reviewed here, especially since excellent treatments are available elsewhere [130, 179–182].

The general idea of using generalized Shanks transformations to accelerate the convergence of series is equivalent to evaluating a sequence of Padé approximants to the series Eq. 260 at $x = 1$. In practice, it is not necessary to explicitly construct the Padé approximants and evaluate them at $x = 1$. Instead of this cumbersome procedure, the values of the Padé approximants can be calculated very efficiently directly from the coefficients of the series using the scalar epsilon algorithm [180–182].

It is also of some interest that the LA and CG algorithm are closely related to a vector version of the epsilon algorithm [178]. The vector epsilon algorithm can even be used to diagonalize matrices [183]! Matrix-valued continued fractions have also been applied to the solution of Fokker–Planck operators [119] and boundary problems [184].

### 4. *Summary*

The mathematical thread that ties these various ideas together is the determinants that arise throughout in the discussions about the existence of the various approximations. The determinants $\Delta_m$ play a crucial role in:

- The existence of a well-behaved complex symmetric LA (Appendix A, Eqs. 222 and 223);
- The definition of the denominators of the rational approximants and continued fractions derived from the Lanczos tridiagonal matrix (Eqs. 248 and 250);
- The existence of a $J$-fraction solution of the generalized or relaxation moment problem (Eq. 257); and
- The definition of the relevant class of Padé approximants to $zJ^{(N)}(z)$ (Eq. 259).



A study of the implications of this connection between these different approaches shows that the LA, relaxation moment problem, and Padé approximant methods are all essentially equivalent from an analytic point of view. The difference between these methods lies in the stability and efficiency of the numerical algorithms required to implement them on a computer.

In his thesis, Dammers [49] compared the Padé approximant [46, 49], relaxation moment [47, 48], and Lanczos methods [3] with respect to the invariance of the calculated spectrum under origin shifts. It is obvious that

$$J(z) = \mathbf{v}[(z - z_0)\mathbf{I} + (\mathbf{A} + z_0\mathbf{I})]^{-1}\mathbf{v}$$

is independent of the value chosen for $z_0$. However, the exact analytical invariance of the preceding expression is not necessarily reflected by the numerical procedures under comparison. In his study, only the results of the LA were found to be invariant with respect to origin shifts of this type. The other two methods show substantial deviations in the computed value of $J(z)$ as a function of $z_0$.

Dammers analyzed this phenomenon in terms of the problems associated with the iterative generation of the sequence of moments in the Padé and relaxation moment methods. The LA does not suffer from such problems since the Lanczos projections automatically cancel the effect of the origin shift to the machine precision at every step. These observations are consistent with the well-known problems associated with the numerical stability of calculating high-order Padé approximants [181] and the need for extended precision arithmetic when performing very slow-motional calculations with the relaxation moment method [47, 48]. The use of extended precision arithmetic has never been necessary in any of our calculations with the LA, even for two-dimensional ESE spectra [4, 30, 88] and extremely complicated slow-motional CWESR studies [20, 21]. These studies indicate that the LA is the method of choice for calculating slow-motional ESR spectra.

Whitehead and Watt [185, 186] have also examined this matter from an analytical view point. They found that the rearrangement to the calculation of the elements of the continued fraction from the moments in a manner that eliminates terms that would cancel in exact arithmetic leads to the LA in a natural way. These ideas are also related to the cumulant and moment

expansion methods that have previously been applied in analytical studies of magnetic resonance and relaxation [10, 11, 175, 187–189].

The penultimate point of this appendix has been to summarize the various mathematical methods that have been developed to calculate spectral functions. In his pioneering work, Mori [107] presented a general and powerful statistical mechanical methodology to develop the continued fraction form of spectral functions. These different mathematical methods have been shown (e.g. in this appendix) to be analytically equivalent to one another and ultimately are equivalent to the Mori method as well [43, 44, 49, 111, 112]. However, these different mathematical realizations give rise to numerical algorithms which differ significantly in their numerical reliability and stability.

### C. Bilinearly Metric Spaces and Relaxation Phenomena

In a very general manner, Moro and Freed [3] showed that calculations of magnetic resonance spectra and related Fokker–Planck forms can be simplified by taking advantage of the fact that the matrix of the operator generating the time evolution of the system can always be put into a complex symmetric form. If $\mathbf{A}$ is the non-Hermitian complex symmetric matrix generating the time evolution, and $\mathbf{A}^\dagger$ is its complex-conjugate transpose, the following holds:

- The eigenvectors of $\mathbf{A}$ are not orthogonal with respect to the usual unitary space inner product.
- The eigenvectors of $\mathbf{A}$ and $\mathbf{A}^\dagger$ are not identical; instead, they are related by complex conjugation.
- The time evolution of the system is not unitary but has a different structure related to a complex orthogonal transformation.
- The transformation (if it exists) that diagonalizes the matrix that generates the time evolution of the system is also complex orthogonal, not unitary.

It is obvious that there are many aspects of the structure of complex vector spaces equipped with the usual unitary space norm and scalar product that make theoretical and numerical studies quite cumbersome for this class of problems. Two separate approaches have been developed to cope with these difficulties. One approach is to retain the usual unitary space inner-product structure and introduce two basis sets [111, 112]. The basis vectors within each set are not orthogonal to one other but are orthogonal to all but one member of the other set. This type of basis is called biorthogonal. The second approach is to simply redefine the norm and scalar product and retain a single basis set. Though it is not apparent, these two approaches are actually identical for complex symmetric matrices [4, 44]. It is in the spirit of the second approach that this section is directed toward the systematic development of the generalization of the concepts of the norm and scalar product in a finite dimensional complex vector space in which complex symmetric and complex orthogonal matrices play the same respective roles as Hermitian and unitary matrices play in the usual scalar product space. In the latter part of the section some of the peculiar aspects of geometry in this new space will be discussed. All of this material is well known to mathematicians but has not been presented in this fashion to the chemical physics community.



The approach used here is that based on the work of Heuvers [190] on the types of functions that can serve as scalar products in complex vector spaces. In this manner any explicit definition of dual spaces and related matters can be avoided. As emphasized by Znojil [191], it is unnecessary to define a scalar product of any type in order to use the LA. The Lanczos recursion relation given by Eq. 212 can be simply viewed as a linear three-term recursion relation that is well defined by the algebraic properties of the complex vector space in which it is defined. The middle-of-the-road view taken here is valuable in that some notion of geometry can be defined through the introduction of a new type of scalar product that is more natural for the problems at hand. This new geometry is naturally associated with the known mathematical structure of the class of problems under consideration (see Section VI.B).

The concept of linear independence of a set of complex $N$ vectors is strictly a property of the algebraic structure associated with the complex vector space and thus is unrelated to the definitions of orthogonality or normalization of vectors.

***Definition 10 (Linear Independence)*** *A set of vectors* $\mathbf{y}_i$, $\mathbf{y}_i \in \mathbf{V} \subseteq \mathscr{C}^N$, $i = 1, 2, \ldots, M$, *is called linearly dependent if there exists a corresponding set of constants* $\alpha_i \in \mathscr{C}$, $i = 1, \ldots, M$, *not all of which are zero, such that*

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_M \mathbf{x}_M = 0. \tag{261}$$

*A set of vectors that is not linearly dependent is called linearly independent. The linear independence of a set of vectors implies that Eq. 261 can only be satisfied if all the constants are simultaneously equal to zero.*

***Definition 11 (Bases and Dimension)*** *A finite set of vectors* $\mathbf{f}_i \in \mathbf{V}$, $\mathbf{i} = 1, \ldots, M$, *is called a basis of* $\mathbf{V}$ *if the set* $\mathbf{f}_i$ *is linearly independent and if for any* $\mathbf{x} \in \mathbf{V}$ *there exist constants* $x_i \in \mathscr{C}$, $i = 1, 2, \ldots, M$, *such that*

$$\mathbf{x} = \sum_{i=1}^{M} x_i \mathbf{f}_i. \tag{262}$$

*It can be shown that the maximum number of linearly independent vectors in* $\mathbf{V}$ *is equal to* $M$. *This number is independent of the choice of basis and is called the dimension of* $\mathbf{V}$.

For reference, it is useful to introduce the concept of bilinear forms as a stepping stone to the usual definitions of norm and inner product. This will enable us to formulate the required generalization of the norm and scalar product of a vector for the class of problems at hand.

***Definition 12 (Primitive Bilinear Form)*** *A complex scalar-valued function* $\rho(\mathbf{x}, \mathbf{y})$ *of two independent vector variables* $\mathbf{x}, \mathbf{y} \in \mathbf{V}$ *is called a primitive bilinear form on* $\mathbf{V} \times \mathbf{V}$ *if, for all vectors* $\mathbf{z} \in \mathbf{V}$ *and scalars* $\alpha, \beta \in \mathscr{C}$,

$$\rho(\mathbf{x} + \mathbf{z}, \mathbf{y}) = \rho(\mathbf{x}, \mathbf{y}) + \rho(\mathbf{z}, \mathbf{y}), \tag{263}$$

$$\rho(\mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z}) = \alpha\rho(\mathbf{x}, \mathbf{y}) + \beta\rho(\mathbf{x}, \mathbf{z}). \tag{264}$$

The adjective *primitive* comes from the fact that the usual types of bilinear forms are special cases of these primitive bilinear forms. It is very important to realize that this definition does not specify the relationship of $\rho(\mathbf{x}, \mathbf{y})$ to $\rho(\mathbf{y}, \mathbf{x})$. The reason that this "zero-order" definition is introduced here is that it will serve as a common starting point for the various types of bilinear metrics. The definition of a particular class of bilinear forms involves the specification of the properties of a primitive bilinear form with respect to interchange of the vector variables.

***Definition 13 (Classes of Bilinear Forms)*** *A primitive bilinear form* $\rho(\mathbf{x}, \mathbf{y})$ *on* $\mathbf{V} \times \mathbf{V}$ *is a symmetric, antisymmetric, or Hermitian bilinear form in* $\mathbf{V}$ *if, for all vectors* $\mathbf{x}, \mathbf{y} \in \mathbf{V}$,

$$\rho(\mathbf{y}, \mathbf{x}) = \rho(\mathbf{x}, \mathbf{y}), \tag{265}$$

$$\rho(\mathbf{y}, \mathbf{x}) = -\rho(\mathbf{x}, \mathbf{y}), \tag{266}$$

$$\rho(\mathbf{y}, \mathbf{x}) = \rho^*(\mathbf{x}, \mathbf{y}), \tag{267}$$

*respectively.*

These three types of bilinear forms generate the three essentially unique types of scalar products in complex vector spaces [190].

***Definition 14 (Scalar Product)*** *A scalar product on* $\mathbf{V}$ *is a bilinear form on* $\mathbf{V} \times \mathbf{V}$ *that satisfies* $\rho(\mathbf{z}, \mathbf{x}) = 0$ *where* $\mathbf{z} \neq 0$, *for all* $\mathbf{x}$ *if and only if* $\mathbf{z} = 0$. *A scalar product is defined to be symmetric, antisymmetric, or Hermitian if the associated bilinear form has this property.*



In this discussion only symmetric and Hermitian scalar products will be considered. A complex vector space with a positive definite Hermitian scalar product is known as a unitary space, whereas if the scalar product is indefinite, it is known as an indefinite inner-product space. A complex vector space with a symmetric scalar product is called a complex orthogonal, or complex Euclidean, space.

The motivation for introduction of a scalar product in a complex vector space is that it can be used to define a geometry on the space.

***Definition 15 (Bilinearly Metric Space)*** *A complex vector space* $\mathbf{V}$ *with a scalar product is a bilinearly metric space.*

The two most important geometric concepts for the present discussion are those of orthogonality and normalization.

***Definition 16 (Orthogonality of Vectors)*** *Two vectors* $\mathbf{x}, \mathbf{y} \in \mathbf{V}$ *are orthogonal with respect to the scalar product* $\rho(\cdot, \cdot)$ *if* $\rho(\mathbf{x}, \mathbf{y}) = 0$.

***Definition 17 (Norms and Normalizable Vectors)*** *A vector* $\mathbf{x} \in \mathbf{V}$ *is normalizable with respect to the scalar product* $\rho(\cdot, \cdot)$ *if there exists a constant* $\beta \in \mathscr{C}$ *such that*

$$\rho(\beta\mathbf{x}, \beta\mathbf{x}) = 1. \tag{268}$$

*The constant* $\beta$ *is called the norm of* $\mathbf{x}$ *if some convention is specified to uniquely determine it. If a vector is nonzero but not normalizable, it will be called nonnormalizable or quasi-null, and it cannot be assigned a norm.*

From this definition it is easy to see that all vectors such that $\rho(\mathbf{x}, \mathbf{x}) \neq 0$ are normalizable. In particular, the usual scalar product of two complex vectors $\mathbf{x}, \mathbf{y} \in \mathscr{C}^N$, $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{N} x_i^* y_i$ and the vector norm $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{N} |x|^2}$ conform to the definitions given here of a positive definite Hermitian scalar product.

The adjoint of a matrix with respect to a scalar product is also an important quantity in the study of bilinearly metric spaces.

***Definition 18 (Adjoint with Respect to Scalar Product)*** *The adjoint* $\mathbf{A}^\ddagger$ *of a square complex matrix* $\mathbf{A}$ *with respect to a scalar product* $\rho(\mathbf{x}, \mathbf{y})$ *is defined by*

$$\rho(\mathbf{x}, \mathbf{A}\mathbf{y}) = \rho(\mathbf{A}^\ddagger\mathbf{x}, \mathbf{y}). \tag{269}$$

*A matrix* $\mathbf{A}$ *is called self-adjoint with respect to* $\rho(\cdot, \cdot)$ *if* $\mathbf{A} = \mathbf{A}^\ddagger$.

If the scalar product is Hermitian, $\mathbf{A}^{\ddagger} = \mathbf{A}^{\dagger}$; if it is symmetric, $\mathbf{A}^{\ddagger} = \mathbf{A}^{tr}$. A closely related construct is that of an isometric matrix or transformation in a bilinearly metric space.

***Definition 19 (Isometric Transformation)*** *An isometric transformation B on a complex vector space* **V** *with scalar product* $\rho(\mathbf{x}, \mathbf{y})$ *is a linear transformation that preserves the scalar product between any two vectors* $\mathbf{x}, \mathbf{y} \in \mathbf{V}$, *that is,*

$$\rho(\mathbf{x}, \mathbf{y}) = \rho(B\mathbf{x}, B\mathbf{y}). \tag{270}$$

The matrix of an isometric transformation will also be called isometric.

A table summarizing the features of real Euclidean, unitary, and complex orthogonal spaces is given in the following.

| Space | Field | Symmetry of Scalar Product | Isometric Transformation |
|---|---|---|---|
| Euclidean | $\mathscr{R}$ | Symmetric | Real orthogonal |
| Unitary | $\mathscr{C}$ | Hermitian | Unitary |
| Complex orthogonal | $\mathscr{C}$ | Symmetric | Complex orthogonal |

The geometry in a unitary space is a natural generalization of the geometry of real Euclidean spaces to complex vector spaces (e.g., all nonzero vectors have positive norms, one can define an angle between any two nonzero vectors, the vector norm satisfies the triangle inequality, etc.). The infinite dimensional analogue of a unitary space is a Hilbert space. Despite the wide use of unitary and Hilbert spaces in modern physical theories, there are many applications where these definitions are too restrictive. One way to generalize these ideas that has proven quite useful is to relax the requirement that the inner product be positive definite and work in an indefinite inner-product space rather than a unitary space. The general mathematical theory for both finite and infinite dimensional indefinite inner-product spaces has been worked out (192–198). These ideas have been applied to the study of quantum field theory [199], electrical networks [200], and the stability of systems of differential equations with periodic coefficients [201]. The theory of indefinite inner-product spaces has been tailored to the study of problems where the time evolution of the system is governed by an operator **B** that is self-adjoint with respect to an indefinite Hermitian scalar product. It can be shown that a general complex matrix **B** is similar to its Hermitian conjugate if and only if there exists a nonsingular Hermitian matrix **J** such that $\mathbf{BJ} = \mathbf{JB}^{\dagger}$ [91, 94]. This Hermitian matrix can then be used to define a Hermitian bilinear form in which **B** is self-adjoint (i.e., $\rho(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{N} J_{ij} x_i^* y_j$, where $J_{ij} = J_{ij}^*$ and $\mathbf{BJ} = \mathbf{JB}^{\dagger}$ so that $\rho(\mathbf{x}, \mathbf{By}) = \rho(\mathbf{Bx}, \mathbf{y})$). It is not always possible to find a scalar product with these properties if the time evolution is governed by a complex symmetric matrix. In fact, the existence of the matrix **J** with the preceding properties implies that *B* is similar to a matrix with only real elements [87, 91]. This is certainly not the case for all complex symmetric matrices since any general square complex matrix is similar to a complex symmetric matrix [86, 87]. If the matrix governing the time evolution of the system is not similar to its Hermitian conjugate, the desired Hermitian scalar product does not exist. In practice, since there are no rules or algorithms to derive the matrix **J** (if it exists) from the matrix **B**, the indefinite inner-product formalism cannot be used unless the equation for the time evolution of the system or other physical insight can be used to specify **J**. It appears that the indefinite inner-product spaces are not the most convenient framework for the study of magnetic resonance lineshapes and related phenomena since the construction of the required indefinite inner product is difficult or impossible.



An alternative to the loosening of the requirement that the (Hermitian) inner product be positive definite is to use a symmetric rather than a Hermitian bilinear form in the definition of the fundamental scalar product. In this manner, it is always possible to construct a scalar product in which the complex symmetric matrix in question is self-adjoint. The simplest choice for a symmetric scalar product is to use the $N \times N$ identity matrix to generate a symmetric bilinear form $\sigma(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{N} x_i y_i$ in the same way as is done to construct the usual inner product $\rho(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{N} x_i^* y_i$. It is important to note that all complex symmetric matrices are self-adjoint with respect to the scalar product $(\cdot, \cdot)$ the same way that all Hermitian matrices are self-adjoint with respect to $\langle \cdot, \cdot \rangle$.

The geometry in complex orthogonal spaces is much more complicated than that of unitary spaces. The description of the geometry in complex orthogonal spaces given here relies heavily on the work of Choudhury and Horn [90] and Malcev [128]. One of the biggest differences is the fact that the orthogonality of two vectors in a complex orthogonal space does not imply that they are linearly independent. A simple example of this phenomenon is given by the two vectors $\mathbf{u} = (1, i)^{tr}$ and $\mathbf{v} = \alpha(1, i)$ in $\mathscr{C}^2$ equipped with the symmetric scalar product $(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2$. It is easy to verify that $(\mathbf{u}, \mathbf{v}) = 0$ for all $\alpha \in \mathscr{C}$. Note that when $\alpha = 1$ this example also demonstrates there can be nonzero vectors that are not normalizable. The analogue in a complex orthogonal space of a set of orthonormal vectors in a unitary space is a set of rectanormal vectors (90).

***Definition 20 (Rectanormality of a Set of Vectors)*** *A set of vectors* $\mathbf{y}_i \in \mathbf{V}$, $i = 1, 2, \ldots, M$, *is called retanormal with respect to the symmetric scalar product* $\sigma(\cdot, \cdot)$ *if*

$$\sigma(\mathbf{y}_i, \mathbf{y}_j) = \delta_{i,j} \quad \textit{for all } i, j. \tag{271}$$

Obviously a rectanormal set of vectors cannot contain any non-normalizable vectors.

Given the fact that not all nonzero vectors in a complex orthogonal space are normalizable, the existence of a well-defined complex orthogonal space analogue of the unitary space Gram–Schmidt orthonormalization procedure is called into question. Choudhury and Horn have studied this problem and have proven under what conditions a complex orthogonal space Gram–Schmidt rectanormalization procedure can be established. Before stating these conditions, it is useful to define the triangular equivalence of two sets of vectors.

***Definition 21 (Triangular Equivalence)*** *Two sets of vectors* $\mathbf{x}_i$, $i = 1, 2, \ldots, M$, *and* $\mathbf{y}_j$, $j = 1, 2, \ldots, M$, *from an N-dimensional complex orthogonal space* $\mathbf{V}$ *are called triangularly equivalent if the span of the set* $\mathbf{x}_i$, $i = 1, 2, \ldots, K$, *is identical with the span of* $\mathbf{y}_j$, $j = 1, 2, \ldots, K$, *for all* $K \leqslant M$.

The Gram–Schmidt rectanormalization procedure is the same as the Gram–Schmidt orthonormalization procedure except that the unitary space Hermitian scalar product is replaced everywhere by the symmetric scalar product appropriate for the complex orthogonal space (see Appendix A). The conditions under which this procedure makes sense is the topic of the following theorem by Choudhury and Horn [90]:

***Theorem 2 (Existence of Gram–Schmidt Rectanormalization Procedure)*** *Given a set of linearly independent vectors* $\mathbf{x}_i$, $i = 1, 2, \ldots, M$, *in a complex orthogonal space, define the sequence of matrices* $X_1, X_2, \ldots, X_M$, *where the elements of the* $j$*th column of* $X_i$ *are given by the elements of* $\mathbf{x}_j$ *for* $j \leqslant i$. *The set* $\mathbf{x}_i$ *is triangularly equivalent to a rectanormal set* $\mathbf{y}_j$ (*cf. Def. 21*) *if and only if*

$$\det(X_i^{\mathrm{tr}} X_i) \neq 0 \quad \text{for } 1 \leqslant i \leqslant M. \tag{272}$$

The Gram–Schmidt rectanormalization of the sequence of Krylov vectors generated by a complex symmetric matrix and an arbitrary normalizable starting vector is discussed in Appendix A. The existence of the Gram–Schmidt rectanormalization procedure for the Krylov vectors implies that the complex symmetric LA is well defined. This result was referred to in Appendix A.

In summary, complex orthogonal spaces enjoy the following properties:

- All CSMs are self-adjoint.
- The spectral function can be expressed as a symmetric bilinear form or a sum of symmetric bilinear forms.



- If two CSMs are similar, they are related by an isometric transformation. In particular, the transformation that diagonalizes a diagonalizable CSM is isometric.
- Under suitable restrictions, a Gram–Schmidt rectanormalization procedure can be used to sequentially convert a linearly independent set of vectors into a triangularly equivalent rectanormal set. The conditions under which this procedure is valid are the same as those needed to have a well-defined complex symmetric LA.
- The rectanormality of vectors is closely related to the known properties of the eigenvectors of general Fokker–Planck operators [119] (see Section VI.B).

Thus, complex orthogonal spaces can be used as a similar type of framework for the calculation of magnetic resonance spectra and spectral densities derived from Fokker–Planck forms as unitary spaces provide for quantum-mechanical problems in the absence of relaxation.

### D. Matrix Elements in ESR Problems

We consider a spin probe with one hyperfine interaction, described by the spin Hamiltonian (in frequency units) in the high-field approximation

$$\hat{\mathscr{H}}(\Omega) = \omega_0 \hat{S}_z + \gamma_e a \hat{S}_z \hat{I}_z + \sum_{K,M,\mu} (-1)^K F_\mu^{(2,-K)} \mathscr{D}_{KM}^2(\Omega) A_\mu^{(2,M)}, \tag{273}$$

where the nuclear Zeeman interaction is not taken into account. We assume that the rotational motion is described by the symmetrized Smoluchowski equation:

$$\tilde{\mathbf{\Gamma}} = [\mathbf{J} - (1/2k_B T)(\mathbf{J}V)]\mathbf{R}[\mathbf{J} + (1/2k_B T)(\mathbf{J}V)], \tag{274}$$

where $\mathbf{J}$ is the generator of infinitesimal rotations of the molecule, $\mathbf{R}$ is the diffusion tensor, and $V$ is the pseudopotential for oriented phases.

We provide only the simplified expression for the matrix elements of $\tilde{\mathbf{\Gamma}} - i\mathbf{L}$ resulting from the following conditions: (i) the contribution of the nonsecular terms to the spectrum is negligible, (ii) the magnetic tensors ($g$ and $A$) and $\mathbf{R}$ tensor are all diagonal in the same molecular frame, (iii) the diffusion tensor is axially symmetric with respect to the $z$ axis of the molecular frame, and (iv) the orienting pseudopotential is axially symmetric and is given by the relation

$$-V(\Omega)/k_B T = \mathscr{D}_{00}^2(\Omega). \tag{275}$$

The matrix elements of the operator $\tilde{\mathbf{\Gamma}} - i\mathbf{L}$ will be calculated in the direct-product space of generalized spherical harmonics and the space of the spin transitions $|m', m''\rangle$, where $m'$ and $m''$ are the eigenvalues of $\hat{I}_z$ of the initial and

final states, respectively, and the electron spin projection indices are implicit,

$$|L, M, K; m', m''\rangle = \left(\frac{2L+1}{8\pi^2}\right)^{1/2} \mathscr{D}^L_{MK}(\Omega)|m', m''\rangle, \tag{276}$$

where $L = 0, 1, 2, \ldots, L_{\max}$, $K = 0, \pm 2, \pm 4, \ldots, \min(L, K_{\max})$, and $M = 0, \pm 1, \pm 2, \ldots, \min(L, M_{\max})$.

In general, the symmetry of the operator and the definition of the starting vector allow one to work with a reduced basis set. The symmetries in the molecular frame allow one to redefine basis elements as

$$|L, M, K'; m', m''\rangle = [2(1+\delta_{0,K})]^{-1/2}|L, M, K; m', m''\rangle + (-1)^L|L, M, -K; m', m''\rangle), \tag{277}$$

with $K'$ nonnegative and even. The symmetry in the laboratory frame allows a further reduction to redefined basis elements:

$$|L, M', K'; p, q\rangle = [2(1+\delta_{0,M}\delta_{0,p})]^{-1/2}(|L, M, K'; p, q\rangle + (-1)^{L+M}|L, -M, K'; -p, q\rangle), \tag{278}$$

where the index $M'$ is now nonnegative. The quantum numbers $p$ and $q$ are defined as $p = m' - m''$ and $q = m' + m''$, where $p = 0$ ($p \neq 0$) for an allowed (forbidden) ESR transition. Our calculations were performed specifically for $^{14}$N-containing nitroxides with nuclear spin $I = 1$, so $m', m'' = 0, \pm 1$. Thus, $p = 0, \pm 1, \pm 2$ and $q = -Q, -Q+2, \ldots, Q-2, Q$, where $Q = 2 - |p|$. In isotropic fluids and in ordered fluids where the symmetry axis of the potential is colinear with the static magnetic field, there exist symmetries such that only the terms with $p = M'$ need be considered so that $M' = 0, 1, 2$. In the following formulas, the primes on $M$ and $K$ are dropped for convenience.

The matrix elements for the diffusion operator in this basis are given by

$$\begin{aligned}
&\langle L_1, M_1, K_1; p_1, q_1|\tilde{\Gamma}|L_2, M_2, K_2; p_2, q_2\rangle \\
&= \langle p_1, q_1|p_2, q_2\rangle \delta_{K_1,K_2}\delta_{M_1,M_2} R_\perp \\
&\times \left\{ \left[ L_1(L_1+1) + \left(\frac{R_\parallel}{R_\perp} - 1\right) K_1^2 + \tfrac{3}{10}(\varepsilon_0^2)^2 \right] \delta_{L_1,L_2} + (-1)^{M_1+K_1} N_L(L_1, L_2) \right. \\
&\times \left[ 3\varepsilon_0^2\left(1 - \tfrac{1}{14}\varepsilon_0^2\right) \begin{pmatrix} L_1 & 2 & L_2 \\ K_1 & 0 & -K_2 \end{pmatrix} \begin{pmatrix} L_1 & 2 & L_2 \\ M_1 & 0 & -M_2 \end{pmatrix} \right. \\
&\left. \left. + \tfrac{18}{35}(\varepsilon_0^2)^2 \begin{pmatrix} L_1 & 4 & L_2 \\ K_1 & 0 & -K_2 \end{pmatrix} \begin{pmatrix} L_1 & 4 & L_2 \\ M_1 & 0 & -M_2 \end{pmatrix} \right] \right\},
\end{aligned} \tag{279}$$



where $N_L(L_1, L_2) = [(2L_1+1)(2L_2+1)]^{1/2}$. The components of the starting vector $|v\rangle$ are given as

$$\langle L, M, K; p, q|v\rangle = \delta_{0,K}\delta_{0,M}\langle p, q|S_-\rangle \int d\Omega \mathscr{D}^L_{00}(\Omega) P_0^{1/2}(\Omega). \tag{280}$$

Finally, the Liouville operator has the following matrix elements:

$$\begin{aligned}
&\langle L_1, M_1, K_1; p_1, q_1|\mathbf{L}|L_2, M_2, K_2; p_2, q_2\rangle \\
&= \langle L_1, M_1, K_1; q_1|L_2, M_2, K_2; q_2\rangle(\omega_0 + \tfrac{1}{2}aq_1\gamma_e) + (-1)^{M_1} \\
&\times [(1+\delta_{2,K_1+K_2})(1+\delta_{1,M_1+M_2})]^{1/2} N_L(L_1, L_2) \\
&\times \begin{pmatrix} L_1 & 2 & L_2 \\ K_1 & K_2 - K_1 & K_2 \end{pmatrix} \begin{pmatrix} L_1 & 2 & L_2 \\ M_1 & M_2 - M_1 & M_2 \end{pmatrix} \\
&\times \sum_\mu F_\mu^{(2, K_2 - K_1)} G_\mu(M_1, q_1; M_2, q_2),
\end{aligned} \tag{281}$$

$$G_g(M_1, q_1; M_2, q_2) = \frac{\sqrt{2/3}\,\omega_0}{\mathrm{Tr}\{g\}}, \tag{282}$$

$$\begin{aligned}
G_A(M_1, q_1; M_2, q_2) &= \delta_{M_1,M_2}\delta_{q_1,q_2} q_1\gamma_e/\sqrt{6} + \delta_{1,|M_1 - M_2|}\delta_{q_1,q_2\pm 1}\gamma_e(M_2 - M_1) \\
&\times \{I(I+1) - \tfrac{1}{4}(q_1 \pm M_1[M_1 - M_2])(q_2 \pm M_2[M_1 - M_2])\}^{1/2}.
\end{aligned} \tag{283}$$

We remark that the matrix associated with $\tilde{\mathbf{\Gamma}} - i\mathbf{L}$ is complex symmetric, both $\tilde{\mathbf{\Gamma}}$ and $\mathbf{L}$ being real and symmetric.

While it is a sparse matrix, the structure of the matrix elements is clearly complicated. The more general case, wherein the preceding simplifying assumptions are removed, leads to an even more complicated (but still sparse) matrix structure, as is detailed elsewhere [20].

## References


1. C. Lanczos, *J. Nat. Bur. Stand.* **45**, 255 (1950).
2. M. R. Hestenes and E. Stiefel, *J. Nat. Bur. Stand.* **49**, 409 (1952).
3. G. Moro and J. H. Freed, *J. Chem. Phys.* **74**, 3757 (1981).
4. K. V. Vasavada, D. J. Schneider, and J. H. Freed, *J. Chem. Phys.* **86**, 647 (1987).
5. J. H. Freed, G. V. Bruno, and C. F. Polnaszek, *J. Phys. Chem.* **75**, 3385 (1971).
6. J. H. Freed, G. V. Bruno, and C. Polnaszek, *J. Chem. Phys.* **55**, 5270 (1971).
7. J. H. Freed, in *Electron-Spin Relaxation in Liquids*, L. T. Muus and P. W. Atkins, Eds., Plenum, New York, 1972.
8. J. H. Freed, *Ann. Rev. Phys. Chem.* **23**, 265 (1972).

9. C. F. Polnaszek, G. V. Bruno, and J. H. Freed, *J. Chem. Phys.* **58**, 3185 (1973).
10. J. H. Freed, *J. Chem. Phys.* **49**, 376 (1968).
11. B. Yoon, J. M. Deutch, and J. H. Freed, *J. Chem. Phys.* **62**, 4687 (1975).
12. J. H. Freed, in *Spin Labeling: Theory and Applications*, Vol. I, L. Berliner, Ed., Academic New York, 1976, Chapter 3.
13. S. A. Goldman, G. V. Bruno, C. F. Polnaszek, and J. H. Freed, *J. Chem. Phys.* **56**, 716 (1972).
14. J. S. Hwang, R. P. Mason, L. P. Hwang, and J. H. Freed, *J. Phys. Chem.* **79**, 489 (1975).
15. G. V. Bruno and J. H. Freed, *J. Phys. Chem.* **78**, 935 (1974).
16. C. F. Polnaszek and J. H. Freed, *J. Phys. Chem.* **79**, 2283 (1975).
17. R. F. Campbell and J. H. Freed, *J. Phys. Chem.* **84**, 2668 (1980).
18. E. Meirovitch and J. H. Freed, *J. Phys. Chem.* **84**, 2459 (1980).
19. M. Shiotani, G. Moro, and J. H. Freed, *J. Chem. Phys.* **74**, 2616 (1981).
20. E. Meirovitch, D. Igner, E. Igner, G. Moro, and J. H. Freed, *J. Chem. Phys.* **77**, 3915 (1982).
21. E. Meirovitch and J. H. Freed, *J. Phys. Chem.* **88**, 4995 (1984).
22. E. Meirovitch, A. Nayeem, and J. H. Freed, *J. Phys. Chem.* **88**, 3454 (1984).
23. H. Tanaka and J. H. Freed, *J. Phys. Chem.* **88**, 6633 (1984).
24. H. Tanaka and J. H. Freed, *J. Phys. Chem.* **89**, 350 (1985).
25. L. Kar, E. Ney-Igner, and J. H. Freed, *Biophys. J.* **48**, 569 (1985).
26. A. E. Stillman, L. J. Schwartz, and J. H. Freed, *J. Chem. Phys.* **73**, 3502 (1980).
27. A. E. Stillman, L. J. Schwartz, and J. H. Freed, *J. Chem. Phys.* **76**, 5658 (1982).
28. L. J. Schwartz, A. E. Stillman, and J. H. Freed, *J. Chem. Phys.* **77**, 5410 (1982).
29. J. P. Hornak and J. H. Freed, *Chem. Phys. Lett.* **101**, 115 (1983).
30. G. L. Millhauser and J. H. Freed, *J. Chem. Phys.* **81**, 37 (1984).
31. L. Kar, G. L. Millhauser, and J. H. Freed, *J. Phys. Chem.* **88**, 3951 (1984).
32. G. L. Millhauser and J. H. Freed, *J. Chem. Phys.* **85**, 63 (1986).
33. L. J. Schwartz, G. L. Millhauser, and J. H. Freed, *Chem. Phys. Lett.* **127**, 60 (1986).
34. J. Gorcester and J. H. Freed, *J. Chem. Phys.* **85**, 5375 (1986); ibid. **88**, 4678 (1988).
35. G. L. Millhauser, J. Gorcester, and J. H. Freed, in *Electron Magnetic Resonance of the Solid State*, J. A. Weil, Ed. Canadian Chemical Society, Ottawa (1987), p. 571.
36. R. F. Campbell, E. Meirovitch, and J. H. Freed, *J. Phys. Chem.* **83**, 525 (1979).
37. O. Pschorn and H. W. Spiess, *J. Magn. Reson.* **39**, 217 (1980).
38. E. Meirovitch and J. H. Freed, *Chem. Phys. Lett.* **64**, 311 (1979).
39. C. F. Polnaszek et al., *Biochemistry* **15**, 954 (1976).
40. S. Alexander, A. Baram, and Z. Luz, *Mol. Phys.* **27**, 441 (1974).
41. A. E. Stillman and J. H. Freed, *J. Chem. Phys.* **72**, 550 (1980).
42. G. Moro and J. H. Freed, *J. Phys. Chem.* **84**, 2837 (1980).
43. G. Moro and J. H. Freed, *J. Chem. Phys.* **75**, 3157 (1981).
44. G. Moro and J. H. Freed, in *Large-Scale Eigenvalue Problems*, Mathematical Studies Series, Vol. 127, J. Cullum and R. Willoughby, Eds., Elsevier, NY 1986.
45. R. G. Gordon and T. Messenger, in *Electron Spin Relaxation in Liquids*, L. T. Muus and P. W. Atkins, Eds., Plenum, New York, 1972, Chapter 13.
46. A. J. Dammers, Y. K. Levine, and J. A. Tjon, *Chem. Phys. Lett.* **88**, 192 (1982).
47. M. Giordano, P. Grigolini, D. Leporini, and P. Martin, *Phys. Rev. A* **28**, 2474 (1983).
48. M. Giordano, P. Grigolini, D. Leporini, and P. Martin, *Adv. Chem. Phys.* **62**, 321 (1985).
49. A. J. Dammers, "Numerical Simulation of Electron Spin Resonance Spectra in the Slow Motion Regime," Ph.D. Thesis, Utrecht, 1985.
50. R. Kubo, *Adv. Chem. Phys.* **16**, 101 (1969).
51. A. Abragam, *Principles of Nuclear Magnetism*, Oxford University Press, New York, 1961.
52. J. H. Freed and G. K. Fraenkel, *J. Chem. Phys.* **39**, 326 (1963).
53. J. H. Freed, *J. Chem. Phys.* **41**, 2077 (1964).
54. L. D. Favro, in *Fluctuation Phenomena in Solids*, R. E. Burgess, Ed., Academic, New York, 1965.
55. P. A. Egelstaff, *J. Chem. Phys.* **53**, 2590 (1970).
56. L. C. Biedenharn and J. D. Louck, *Angular Momentum in Quantum Physics*, Addison-Wesley, London, 1981.
57. A. Messiah, *Quantum Mechanics*, Wiley, New York, 1962.
58. M. E. Rose, *Elementary Theory of Angular Momentum*, Wiley, New York, 1957.
59. N. C. Pyper, *Mol. Phys.* **21**, 1 (1971).
60. N. C. Pyper, *Mol. Phys.* **22**, 433 (1971).
61. R. M. Lynden-Bell, in *Electron Spin Relaxation in Liquids*, L. T. Muus and P. W. Atkins, Eds., Plenum, New York, 1972, Chapter 13A.
62. R. M. Lynden-Bell, *Mol. Phys.* **22**, 837 (1972).
63. A. Omont, *Prog. Quant. Electron.* **5**, 69 (1977).
64. V. Gorini, M. Verri, and E. C. G. Sudarshan, in *Group Theoretical Methods in Physics*, K. Wolf, Ed., Lecture Notes in Physics, Vol. 135, Springer-Verlag, New York, 1980.
65. J. B. Pedersen and J. H. Freed, *J. Chem. Phys.* **57**, 1004 (1972).
66. J. B. Pedersen and J. H. Freed, *J. Chem. Phys.* **58**, 2746 (1973).
67. J. B. Pedersen and J. H. Freed, *J. Chem. Phys.* **59**, 2869 (1973).
68. L. P. Hwang and J. H. Freed, *J. Chem. Phys.* **63**, 4017 (1975).
69. J. H. Freed and J. B. Pedersen, *Adv. Mag. Res.* **8**, 1 (1976).
70. J. H. Freed, in *Chemically Induced Magnetic Polarization*, L. Muus et al., Eds., D. Reidel, Dordrecht-Holland, 1977, Chapter 19.
71. K. Wassmer, E. Ohmes, M. Portugall, H. Ringsdorf, and G. Kothe, *J. Am. Chem. Soc.* **107**, 1511 (1985).
72. K. Müller, P. Meier, and G. Kothe, in *Progress in NMR Spectroscopy* (1986).
73. G. P. Zientara and J. H. Freed, *J. Chem. Phys.* **70**, 1359 (1979).
74. G. P. Zientara and J. H. Freed, *J. Chem. Phys.* **71**, 3861 (1979).
75. G. P. Zientara and J. H. Freed, *J. Phys. Chem.* **83**, 3333 (1979).
76. J. P. Korb, M. Ahadi, G. P. Zientara, and J. H. Freed, *J. Chem. Phys.* **86**, 1125 (1987).
77. A. E. Stillman, G. P. Zientara, and J. H. Freed, *J. Chem. Phys.* **71**, 113 (1979).
78. G. P. Zientara and J. H. Freed, *J. Chem. Phys.* **70**, 2587 (1979).
79. G. P. Zientara and J. H. Freed, *J. Chem. Phys.* **71**, 744 (1979).
80. G. H. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1983.
81. C. A. J. Fletcher, *Computational Galerkin Methods*, Springer-Verlag, New York, 1984.
82. J. B. Pederson, *J. Chem. Phys.* **57**, 2680 (1972).

83. M. M. Itzkowitz, *J. Chem. Phys.* **46**, 3048 (1967).
84. A. E. Stillman and R. N. Schwartz, in *Time Domain Electron Spin Resonance*, L. Kevan and R. N. Schwartz, Eds., Wiley-Interscience, New York, 1979, Chapter 5.
85. G. P. Zientara and J. H. Freed, *J. Chem. Phys.* **72**, 1285 (1980).
86. F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1959.
87. R. A. Horn and C. A. Johnson, *Matrix Analysis*, Cambridge University Press, NY 1985.
88. D. J. Schneider, unpublished work, 1986.
89. J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Sparse Eigenvalue Computations*, Birkhäuser, Boston, 1985.
90. D. Choudhury and R. A. Horn, "An Analog of the Gram–Schmidt Algorithm for Complex Symmetric Forms and Diagonalization of Complex Symmetric Matrices," Technical Report No. 454, Department of Mathematical Sciences, Johns Hopkins University, Baltimore, MD, 1986.
91. O. Taussky, *Lin. Alg. Applic.* **5**, 147 (1972).
92. B. D. Craven, *J. Austral. Math. Soc.* **10**, 341 (1969).
93. M. L. Mehta, *Elements of Matrix Theory*, Hindustan, Delhi, India, 1977.
94. J. W. Duke, *Pacific J. Math.* **31**, 321 (1969).
95. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.
96. B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
97a. S. A. Goldman, G. V. Bruno, and J. H. Freed, *J. Chem. Phys.* **59**, 3071 (1973).
97b. S. A. Goldman, Ph.D. Thesis, Cornell University, (1973).
98. L. A. Dalton and L. R. Dalton, in *Multiple Electron Resonance Spectroscopy*, M. Dorio and J. H. Freed, Eds., Plenum, New York, 1979, Chapter 5.
99. J. S. Hyde and L. R. Dalton, in *Spin Labeling: Theory and Applications*, Vol. II, L. Berliner, Ed., Academic, New York, 1976, Chapter 1.
100. N. Benetis, "Nuclear Spin Relaxation in Paramagnetic Metal Complexes and the Slow Motional Problem for the Electron Spin," Ph.D. Thesis, University of Stockholm, 1984.
101. W. Givens, *Nat. Bur. Stand. Appl. Math. Ser.* **29**, 117 (1952).
102. H. Rutishauser, *Proc. Am. Math. Soc. Symp. Appl. Math.* **15**, 219 (1963).
103. P. Anderson and G. Loizou, *J. Inst. Math. Appl.* **12**, 261 (1973).
104. P. Anderson and G. Loizou, *Numer. Math.* **25**, 347 (1976).
105. P. J. Eberlein, *J. Inst. Math. Appl.* **7**, 377 (1971).
106. J. G. Ruffinatti, *Computing* **15**, 275 (1975).
107. H. Mori, *Progr. Theor. Phys.* **33**, 423 (1965).
108. R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II, Nonequilibrium Statistical Mechanics*, Springer Series in Solid-State Science, Vol. 31, Springer-Verlag, New York, 1985.
109. B. J. Berne and R. Pecora, *Dynamic Light Scattering*, Wiley-Interscience, New York, 1976.
110. D. Kivelson and K. Ogan, in *Advances in Magnetic Resonance*, Vol. 7, J. Waugh, Ed., Academic, New York, 1974, Chapter 2.
111. W. A. Wassam, Jr., *J. Chem. Phys.* **82**, 3371 (1985).
112. W. A. Wassam, Jr., *J. Chem. Phys.* **82**, 3386 (1985).
113. J. H. Freed, in *Rotational Dynamics of Small and Macromolecules in Liquids*, T. Dorfmuller and R. Pecora, Eds., Springer-Verlag, New York, (1987), p. 89.
114. J. S. Hwang, K. V. S. Rao, and J. H. Freed, *J. Phys. Chem.* **80**, 1790 (1976).





115. N. G. Van Kampen, *Stochastic Processes in Chemistry and Physics*, North-Holland, Amsterdam, 1981.
116. H. Haken, *Rev. Mod. Phys.* **47**, 67 (1975).
117. H. Haken, *Synergetics*, Springer Series in Synergetics, Vol. 1, 2nd enlarged Ed., Springer-Verlag, New York, 1978.
118. H. Haken, *Advanced Synergetics*, Springer Series in Synergetics, Vol. 20, Springer-Verlag, New York, 1983.
119. H. Risken, *The Fokker–Planck Equation*, Springer Series in Synergetics, Vol. 18, Springer-Verlag, New York, 1984.
120. C. W. Gardiner, *Handbook of Stochastic Methods*, Springer Series in Synergetics, Vol. 13, Springer-Verlag, New York, 1983.
121. P. Hänggi and H. Thomas, *Phys. Rep.* **88**, 207 (1982).
122. H. Risken, *Z. Phy.* **251**, 231 (1972).
123. R. Graham and H. Haken, *Z. Phy.* **243**, 289 (1971).
124. M. Lax, in *Symmetries in Science*, B. Gruber and R. S. Milman, Eds., Plenum, New York, 1980.
125. M. Lax, *Symmetry Principles in Solid State and Molecular Physics*, Wiley, New York, 1974.
126. P. O. Löwdin, *J. Math. Phys.* **24**, 70 (1983).
127. V. V. Voyevodin, *Linear Algebra*, Mir, Moscow, 1983.
128. I. A. Malcev, *Foundations of Linear Algebra*, Freeman, New York, 1963.
129. P. M. Morse and H. Feshbach, *Methods of Mathematical Physics*, 2 Vols., McGraw-Hill, New York, 1953.
130. C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York, 1978.
131. N. G. van Kampen, *Phys. Rep.* **24C**, 171 (1976).
132. U. Frish, in *Probabilistic Methods in Applied Mathematics*, A. T. Bharucha Reid, Ed., Academic, New York, 1968.
133. A. Brissaud and U. Frisch, *J. Math. Phys.* **15**, 524 (1974).
134. J. H. Freed, *J. Chem. Phys.* **66**, 4183 (1977).
135. L. P. Hwang and J. H. Freed, *J. Chem. Phys.* **63**, 118 (1975).
136. A. J. Vega and D. Fiat, *J. Chem. Phys.* **60**, 579 (1974).
137. L. Monchick, *J. Chem. Phys.* **74**, 4519 (1981).
138. W. A. Wassam, Jr. and J. H. Freed, *J. Chem. Phys.* **76**, 6133 (1982).
139. W. A. Wassam, Jr. and J. H. Freed, *J. Chem. Phys.* **76**, 6150 (1982).
140. L. P. Hwang and J. H. Freed, unpublished work.
141. M. Fixman and K. Rider, *J. Chem. Phys.* **51**, 2425 (1969).
142. R. G. Gordon, *J. Chem. Phys.* **44**, 1830 (1966).
143. J. H. Freed, *J. Phys. Chem.* **78**, 1155 (1974).
144. J. H. Freed, in *Multiple Electron Resonance Spectroscopy*, M. Dorio and J. H. Freed, Eds., Plenum, New York, 1979, Chapter 3.
145. J. H. Freed, in *Time Domain Electron Spin Resonance*, L. Kevan and R. N. Schwartz, Eds., Wiley-Interscience, New York, 1979, Chapter 2.
146. L. J. Schwartz, "Molecular Reorientation and Time Domain Electron Spin Resonance," Ph.D. Thesis, Cornell University, 1984.

147. S. A. Goldman, G. V. Bruno, and J. H. Freed, *J. Phys. Chem.* **76**, 1858 (1972).
148. G. V. Bruno, "Application of the Stochastic Liouville Method in Calculating ESR Lin Shapes in the Slow Tumbling Region and ESR-ELDOR Study of Exchange," Ph.D. Thesi Cornell University, 1973.
149. G. V. Bruno and J. H. Freed, *Chem. Phys. Lett.* **25**, 328 (1974).
150. J. H. Freed, *J. Chem. Phys.* **43**, 2312 (1965).
151. J. S. Hyde, J. C. W. Chien, and J. H. Freed, *J. Chem. Phys.* **48**, 4211 (1968).
152. M. D. Smigel, L. A. Dalton, L. R. Dalton, and A. L. Kwiram, *Chem. Phys.* **6**, 183 (1974).
153. H. Barkhuijsen, R. de Beer, W. M. M. J. Bovee, and D. van Ormont, *J. Magn. Reson.* **61**, 46 (1985).
154. R. Kumaresan and D. W. Tufts, *IEEE Trans.* **ASSP-30**, 833 (1982).
155. J. Jeener, B. H. Meier, P. Bachmann, and R. R. Ernst, *J. Chem. Phys.* **71**, 11 (1979).
156. S. A. Dzuba, A. G. Maryasov, K. M. Salikhov, and Yu. D. Tsvetkov, *J. Magn. Reson.* **85**, 9 (1984).
157. U. Eliav and J. H. Freed, *J. Phys. Chem.* **88**, 1277 (1984).
158. K. F. Milfield and N. Moiseyev, *Chem. Phys. Lett.* **130**, 145 (1986).
159. E. Haller, H. Köppel, and L. S. Cederbaum, *J. Mol. Spectrosc.* **111**, 377 (1985).
160. G. Moro, *Chem. Phys.* **106**, 89 (1986).
161. G. Moro and P. L. Nordio, *Chem. Phys. Lett.* **96**, 192 (1983).
162. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York 1964; reprinted by Dover, Mineola, NY, 1975.
163. W. B. Jones and W. J. Thron, *Continued Fractions*, Encyclopedia of Mathematics and It Applications, Vol. 11, Addison-Wesley, Reading, MA, 1980.
164. H. S. Wall, *Analytical Theory of Continued Fractions*, Van Nostrand, New York, 1948 reprinted by Chelsea, New York, 1967.
165. H. S. Wall and M. Wentzel, *Trans. Am. Math. Soc.* **55**, 373 (1944).
166. H. S. Wall and M. Wentzel, *Duke Math. J.* **11**, 89 (1944).
167. H. S. Wall, *Bull. Am. Math. Soc.* **52**, 671 (1946).
168. I. S. Kac and M. G. Krein, *Am. Math. Soc. Trans.* **103**(2), 1 (1974).
169. P. Giannozzi, G. Grosso, and G. Pastori Parravicini, *Phys. Stat. Sol. B* **128**, 643 (1985).
170. G. Grosso and G. Pastori Parravicini, *Adv. Chem. Phys.* **62**, 81 (1985).
171. N. I. Ahiezer and M. G. Krein, *Some Problems in the Theory of Moments*, Transl. Math Monographs, Vol. 2, American Mathematical Society, Providence, RI, 1962.
172. N. I. Ahiezer, *The Classical Moment Problem*, Oliver and Boyd, Edinburgh, 1965.
173. J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, American Mathematical Society, New York, 1943.
174. R. G. Gordon, *J. Math. Phys.* **9**, 1087 (1968).
175. F. Lado, J. D. Parker, and G. W. Memory, *Phys. Rev. B* **4**, 1406 (1971).
176. A. Lonke, *J. Math. Phys.* **12**, 2422 (1971).
177. J. Cullum and R. A. Willoughby in *Large-Scale Eigenvalue Problems*, Mathematical Studies Series, Vol. 127, J. Cullum and R. A. Willoughby, Eds. Elsevier, NY 1986.
178. C. Brezinski, *Padé-Type Approximation and General Orthogonal Polynomials*, International Series of Numerical Mathematics, Vol. 50, Birkhäuser, Boston 1980.
179. G. A. Baker, Jr., *Essentials of Padé Approximants*, Academic, New York, San Francisco, London, 1975.
180. C. Brezinski, *J. Comput. Appl. Math.* **12**, 19 (1985).



181. G. A. Baker, Jr. and P. Graves-Morris, *Padé Approximants*, Encyclopedia of Mathematics and Its Applications, Vols. 13 and 14, Addison-Wesley, Reading, MA, 1983.
182. J. Wimp, *Sequence Transformation Methods and Their Applications*, Academic, New York, 1981.
183. C. Brezinski, *Lin. Alg. Its Applic.* **11**, 7 (1975).
184. F. V. Atkinson, *Discrete and Continuous Boundary Problems*, Academic, New York, 1964.
185. R. R. Whitehead and A. Watt, *J. Phys.* **G4**, 835 (1978).
186. R. R. Whitehead and A. Watt, *J. Phys.* **A14**, 1887 (1981).
187. L. P. Hwang, C. F. Anderson, and H. L. Friedman, *J. Chem. Phys.* **62**, 2098 (1975).
188. J. L. Monroe and H. L. Friedman, *J. Chem. Phys.* **66**, 955 (1977).
189. S. Alexander, A. Baram, and Z. Luz, *J. Chem. Phys.* **61**, 992 (1974).
190. K. J. Heuvers, *Lin. Alg. Its Applic.* **6**, 83 (1973).
191. M. Znojil, *J. Phys. A* **9**, 1 (1976).
192. M. G. Krein, *Am. Math. Soc. Trans.* **93**(2), 103 (1970).
193. M. G. Krein, *Am. Math. Soc. Trans.* **1**(2), 27 (1955).
194. A. Z. Jadczyk, *Rep. Math. Phys.* **2**, 263 (1971).
195. J. Bognár, *Indefinite Inner Product Spaces*, Springer-Verlag, New York, 1974.
196. I. S. Iohvidov, M. G. Krein, and H. Langer, *Introduction to the Spectral Theory of Operators in Spaces with an Indefinite Metric*, Mathematical Research Vol. 9, Akademie-Verlag, Berlin, 1982.
197. I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and Indefinite Scalar Products*, Operator Theory: Advances and Applications, Vol. 8, Birkhäuser, Boston 1983.
198. I. M. Glazman and Ju. I. Ljubič, *Finite-Dimensional Linear Analysis*, MIT Press, Cambridge, MA, 1974.
199. C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, McGraw-Hill, New York, 1980.
200. M. C. Pease III, *Methods of Matrix Algebra*, Mathematics in Science and Engineering, Vol. 16, Academic, New York, 1965.
201. V. A. Yakubovic and V. M. Starzhinskii, *Linear Differential Equations with Periodic Coefficients*, Vols. I and II, Wiley, New York, 1975.
202. G. V. Bruno, J. K. Harrington, and M. P. Eastman, *J. Phys. Chem.* **81**, 1111 (1977).
203. W. J. Lin and J. H. Freed, *J. Phys. Chem.* **83**, 379 (1979).
204. K. V. S. Rao, J. S. Hwang, and J. H. Freed, *Phys. Rev. Lett.* **37**, 515 (1976).
205. S. A. Zager and J. H. Freed, *Chem. Phys. Lett.* **109**, 270 (1984).
206. A. Nayeem, "Electron Spin Relaxation and Molecular Dynamics in Liquid Crystalline Phases," Ph.D. Thesis, Cornell University, 1986.
207. S. A. Zager and J. H. Freed, *J. Chem. Phys.* **77**, 3344, 3360 (1982).
208. G. Moro and P. L. Nordio, *Mol. Cryst. Liq. Crys.* **104**, 361 (1984).
209. G. Moro and P. L. Nordio, *J. Phys. Chem.* **89**, 997 (1985).
210. D. G. Pettifor and D. L. Weaire, Eds., *The Recursion Method and Its Applications*, Springer Series in Solid-State Science, Vol. 58, Springer-Verlag, New York, 1985.
211. I. Efrat and M. Tismenetsky, *IBM J. Res. Develop.* **30**, 184 (1986).
212. E. van der Drift, B. A. C. Rousseeuw, and J. Smidt, *J. Chem. Phys.* **88**, 2275 (1984).
213. Z. Luz and R. Naor, *Mol. Phys.* **46**, 891 (1982).
214. D. Gamliel, Z. Luz, and S. Vega, *J. Chem. Phys.* **85**, 2516 (1986).
215. Added in proof: This has now been successfully implemented.