Derivatives of the Energy Levels of a Molecular System with Respect to Parameters of the Hamiltonian, and Their Application to Magnetic Susceptibility Calculations and Least-squares Fitting

By **Antony B. Blake**, Department of Chemistry, The University of Hull, Hull HU6 7RX

Formulae are given for the first, second, and third partial derivatives, with respect to the parameters a_i, of the eigenvalues of a matrix $\mathscr{H} = \Sigma a_i \mathscr{H}_i$ (where $\mathscr{H}_i$ are Hermitian or real symmetric matrices), in terms of the eigenvalues and eigenvectors of $\mathscr{H}$. The matrix $\mathscr{H}$ may also include bilinear terms of the type $a_i a_j \mathscr{H}_{ij}$. If $\mathscr{H}$ is the Hamiltonian matrix for a molecular system in a magnetic field, so that one of the a_i is the applied field strength and the others represent features of the model such as crystal-field, spin–orbit coupling, and exchange parameters, these results can be used to obtain expressions for the paramagnetic susceptibility which are exact even at very low temperatures, within the limitations of the model and the accuracy with which the eigenvalues and eigenvectors of $\mathscr{H}$ are known. The formulae can also be used to set up least-squares normal equations determining the values of the parameters that give the best fit to experimental susceptibility data (or other observed properties of the system), and to solve those equations iteratively. The formulae are likely to be especially useful whenever it is impossible, because of the complexity of the secular equation, to write explicit expressions for the energy levels and to calculate their derivatives by analytic differentiation, but the generality of the method may make it useful even when analytic differentiation is possible. Examples of its use in the least-squares fitting of magnetic susceptibility data are given.

MAGNETIC susceptibility measurements continue to be widely used to investigate the ground states of transition-metal complexes, especially polynuclear complexes involving exchange interactions, the tendency in recent years being towards the use of very sensitive anisotropic measurement techniques at very low temperatures in order to elucidate the finer details of electronic structure.[1] In interpreting the results of such measurements, the experimenter usually has in mind a theoretical model involving one or more parameters, and wishes to find those values of the parameters that cause the predictions of the model to fit the experimental data as closely as possible, *i.e.* to minimise the weighted sum of the squares of the differences between observed and predicted susceptibilities. Very often the model takes the form of a Hamiltonian $\mathscr{H}$, consisting of a sum of operators, which acts on a chosen set of basis functions to generate a matrix $\mathscr{H}$ from which the properties of the model are to be calculated.

Calculation of the magnetic susceptibility involves taking derivatives of the energy levels with respect to field strength. Provided an algebraic expression for the eigenvalues of $\mathscr{H}$ (including magnetic field effects) can be obtained, this presents no difficulty, but there are many cases of interest in which the secular equation of $\mathscr{H}$ does not factorise sufficiently to permit an analytical solution to the eigenvalue problem, and algebraic expressions for the energy levels then cannot be written. Similarly, calculation of the 'least-squares' values of the parameters (as distinct from their location by an exhaustive search) involves taking derivatives of the susceptibility with respect to the parameters, which again cannot be done algebraically if the secular equation does not factorise. In both cases the usual procedure has been to make approximations to the derivatives (or sometimes, in the fitting problem, to perform a direct search of parameter space), but there are circumstances in which this approach is not wholly satisfactory.

In this paper I discuss the accurate calculation of magnetic susceptibilities from the Hamiltonian matrix, and propose a new approach to the least-squares fitting problem for cases where the secular equation does not factorise adequately. The discussion is based on the observation that, subject to certain conditions, the derivatives of the eigenvalues of a matrix with respect to parameters in which it is linear can be expressed exactly, in closed form, in terms of its eigenvalues and eigenvectors.

CALCULATION OF MAGNETIC SUSCEPTIBILITIES

The magnetic moment operator for any molecular system can be shown to be equal to $-\partial\mathscr{H}/\partial\mathbf{H}$, where $\mathscr{H}$ is the Hamiltonian operator for the system in the presence of an applied magnetic field $\mathbf{H}$.[2] The moment in a state of energy E_n is thus $-\partial E_n/\partial\mathbf{H}$, and the bulk magnetic moment per mole for an assembly of such molecular magnets is given by equation (1), where $\langle\ \rangle_T$ represents the canonical average at temperature T. We shall assume that the sample is magnetically isotropic (extension to anisotropic materials is straightforward), and that $\mathbf{M} = 0$ when $\mathbf{H} = 0$.

$$\mathbf{M}_{mol} = -N_A \langle \partial E/\partial \mathbf{H} \rangle_T = -N_A \left[\frac{\Sigma_n (\partial E_n/\partial \mathbf{H}) \exp(-E_n/\boldsymbol{k}T)}{\Sigma_n \exp(-E_n/\boldsymbol{k}T)} \right] \quad (1)$$

The Faraday, Gouy, and vibrating-sample techniques measure M_{mol} at a known field strength H, and thus yield experimental values of the absolute molar susceptibility $\chi_{mol} = M_{mol}/H$, equation (2). In certain induction tech-

$$\chi_{mol} = -N_A \langle \partial E/\partial H \rangle_T/H \quad (2)$$

niques the quantity measured is the differential susceptibility $\bar{\chi}_{mol} = \partial M_{mol}/\partial H$, equation (3) [obtained by differentiating (1)].[3] It is also possible to measure the sus-

$$\bar{\chi}_{mol} = -N_A[\langle \partial^2 E/\partial H^2 \rangle_T - \langle (\partial E/\partial H)^2 \rangle_T/\boldsymbol{k}T + \langle \partial E/\partial H \rangle_T^2/\boldsymbol{k}T] \quad (3)$$

ceptibility in the limit of zero applied field, $\bar{\chi}_{mol}^0 = \chi_{mol}^0$, which is given by (3) with the third term omitted.

Instead of calculating the derivatives directly, it is common to replace equation (2) by an approximation known as the Van Vleck equation, (5), based on an expansion of E_n as a power series in H, equation (4).[4] Here E_n^0 are the eigenvalues of the zero-field Hamiltonian and $E_n^{(i)}$ are perturbation coefficients for a calculation in which the interaction of the system with the applied magnetic field (the Zeeman term) is taken as a perturbation applied to the eigenstates of the zero-field Hamiltonian. Equation (5) is obtained by neglecting all terms that would make χ

$$E_n = E_n^0 + E_n^{(1)}H + E_n^{(2)}H^2 + \ldots \quad (4)$$

$$\chi_{\mathrm{mol}} = N_{\mathrm{A}}\left(\frac{\Sigma_n\{[E_n^{(1)}]^2/\boldsymbol{k}T - E_n^{(2)}\}\exp(-E_n^0/\boldsymbol{k}T)}{\Sigma_n\exp(-E_n^0/\boldsymbol{k}T)}\right) \quad (5)$$

depend on H (' paramagnetic saturation effects '), thereby implying (*a*) that the perturbation expansion can be terminated at second order, and (*b*) that all energy changes due to the field are small compared with $\boldsymbol{k}T$. When the applied field is vanishingly small, (5) reduces to (3) and is exactly true, but in fields of a few thousand gauss (10^4 G = 1 T), such as are commonly used to measure paramagnetic susceptibilities, (*a*) and (*b*) may not be very good approximations: the first-order energies $E_n^{(1)}H$ may be comparable with some of the zero-field splittings (especially if the Hamiltonian includes weak or competing interactions), and at liquid-helium temperatures they will certainly not be small compared with $\boldsymbol{k}T$.[5a] Thus, in precise, low-temperature work at moderate fields it is desirable to use equation (2) or (3) rather than (5).

If the model is sufficiently simple for the secular equation of $\mathscr{H}$ to factorise with no factor higher than quadratic, the energies E_n can be written as explicit functions of the parameters and of the field strength, allowing the derivatives to be calculated exactly by algebraic methods. Often, however, it is necessary to solve the eigenvalue problem numerically, and this has usually been done by diagonalising the zero-field matrix to obtain the E_n^0, and then calculating the derivatives $\partial E_n/\partial H$ as a truncated perturbation series $E_n^{(1)} + E_n^{(2)}H$ in a basis of zero-field eigenstates, as in the Van Vleck equation.[6] No allowance is thus made for paramagnetic saturation effects. Alternatively, $\partial E_n/\partial H$ has been estimated numerically by diagonalising the Hamiltonian with the Zeeman term included, at two or three closely spaced values of H.[5] (In this method eigenvectors are not needed, but the time saved by omitting their calculation is likely to be used up in repeating the eigenvalue computation.) As we shall see, however, if we take the Hamiltonian matrix for an isolated magnetic system in a magnetic field and obtain its eigenvalues and eigenvectors by computer diagonalisation (in which, of course, it is as easy to include the Zeeman term as not), we can calculate $\partial E_n/\partial H$ and $\partial^2 E_n/\partial H^2$ exactly (limited only by the precision of the computer diagonalisation) and hence obtain the susceptibility χ or $\bar{\chi}$ of the model accurately at any field strength and temperature.

DERIVATIVES OF THE EIGENVALUES OF A HAMILTONIAN

Case A.—We assume that the Hamiltonian $\mathscr{H}$ is linear in a number of scalar parameters a_i, so that in matrix form we have equation (6), where $\mathscr{H}_i$ are known

$$\mathscr{H} = \sum_{i=1}^{r} a_i\mathscr{H}_i \quad (6)$$

matrices calculated in some convenient common basis. Since the derivatives of the energy levels with respect to the a_i are effectively perturbation coefficients for infinitesimal changes in the a_i, they are given by expressions analogous to the formulae of non-degenerate perturbation theory. (The presence of the magnetic field ensures that $\mathscr{H}$ has no degenerate eigenvalues, accidental coincidences apart.) Detailed derivations of the following results are given in Supplementary Publication No. SUP 23004 (7 pp.).*

Let $\mathscr{U}$ be the eigenvector matrix of $\mathscr{H}$, *i.e.* a unitary matrix constructed from the orthonormal column eigenvectors of $\mathscr{H}$, so that $\mathscr{E} = \mathscr{U}^{-1}\mathscr{H}\mathscr{U}$ is a diagonal matrix whose diagonal elements are the eigenvalues E_n of $\mathscr{H}$. Construct the matrices $\mathscr{E}^{(i)} = \mathscr{U}^{-1}\mathscr{H}_i\mathscr{U}$. (These will not in general be diagonal, though they will be Hermitian or real-symmetric.) Then the derivatives $E_{ni}' = \partial E_n/\partial a_i$, $E_{nij}'' = \partial^2 E_n/\partial a_i\partial a_j$, and $E_{nijk}''' = \partial^3 E_n/\partial a_i\partial a_j\partial a_k$ are given by equations (7), (8), and (9), in which the terms with $m = n$ or $m' = n$ are excluded

$$E_{ni}' = \mathscr{E}_{nn}^{(i)} \quad (7)$$

$$E_{nij}'' = 2\sum_{m} \mathrm{Re}[\mathscr{E}_{nm}^{(i)}\mathscr{E}_{mn}^{(j)}]/(E_n - E_m) \quad (8)$$

$$E_{nijk}''' = 2\sum_{m}\sum_{m'} \left[\frac{\sum_{\nu} P_\nu(\mathrm{Re}\{\mathscr{E}_{nm}^{(i)}[\mathscr{E}_{mm'}^{(j)} - \mathscr{E}_{nn}^{(j)}\delta_{mm'}]\mathscr{E}_{m'n}^{(k)}\})}{(E_n - E_m)(E_n - E_{m'})}\right] \quad (9)$$

$$E_n = \sum_{i=1}^{r} a_i E_{ni}' \quad (10)$$

from the sums, and P_ν ($\nu = 1$, 2, or 3) selects one of the three cyclic permutations of the indices ijk. We note also the relationship (10).

Case B.—We shall also need to consider the case where $\mathscr{H}$ includes terms bilinear in certain parameters, as in equation (11), where the matrices $\bar{\mathscr{H}}_i$ are defined in

$$\mathscr{H} = \sum_{i=1}^{r} a_i\mathscr{H}_i + \sum_{i=1}^{r-1}\sum_{j=i+1}^{r} c_{ij}a_ia_j\mathscr{H}_{ij} = \sum_{i=1}^{r} a_i\bar{\mathscr{H}}_i \quad (11)$$

$$\bar{\mathscr{H}}_i = \mathscr{H}_i + \sum_{j=i+1}^{r} c_{ij}a_j\mathscr{H}_{ij} \quad (12)$$

(12), $\mathscr{H}_i$ and $\mathscr{H}_{ij}$ being known matrices and c_{ij} known constants, with $c_{ii} = 0$. (Usually only one of the c_{ij} is non-zero.) If we define $\bar{\mathscr{E}}^{(i)} = \mathscr{U}^{-1}\bar{\mathscr{H}}_i\mathscr{U}$ and $\mathscr{E}^{(ij)} = \mathscr{U}^{-1}\mathscr{H}_{ij}\mathscr{U}$, the derivatives are given by equations (7) to (9), modified by the substitution of $\bar{\mathscr{E}}^{(i)}$ for $\mathscr{E}^{(i)}$ and by inclusion of the extra terms shown in equations (13) and (14). We note also the relationship (15).

$$E_{nij}''' = (\bar{8}) + c_{ij}\mathscr{E}_{nn}^{(ij)} \quad (13)$$

$$E_{nijk}''' = (\bar{9}) + \sum_{\nu} P_\nu\{c_{ij}\sum_{m}\mathrm{Re}[\mathscr{E}_{nm}^{(ij)}\bar{\mathscr{E}}_{mn}^{(k)}]/(E_n - E_m)\} \quad (14)$$

$$\bar{\mathscr{E}}^{(i)} = \bar{\mathscr{E}}^{(i)} + \sum_{j=i+1}^{r} c_{ij}a_j\mathscr{E}^{(ij)} \quad (15)$$

* For details see Notices to Authors No. 7, *J. Chem. Soc., Dalton Trans.*, 1979, Index issue.

APPLICATION TO LEAST-SQUARES FITTING

(*a*) *General.*—Let us suppose that we are measuring some physical property F which depends on the parameters $a_1, a_2, \ldots, a_r$, and that we have a number of observed values S_p of this property measured under conditions (*e.g.* of temperature) such that they correspond to theoretical values F_p. Our aim is to find a set of values of the a_i that minimises the function ρ, equation (16) (w_p being weights assigned to the observations).

$$\rho\ (a_1, a_2, \ldots, a_r) = \sum_p w_p(F_p - S_p)^2 \qquad (16)$$

What I shall call the ' traditional least-squares ' approach starts by equating to zero the first derivative of ρ with respect to each a_i, leading to a set of r simultaneous equations (17), the *normal equations*. Solution of these

$$\sum_p w_p(F_p - S_p)F_{pi}' = 0 \qquad (17)$$

(with suitable precautions to exclude saddle points and local minima of ρ) gives the ' best-fit ' values of the r unknowns a_i. (Note that we can, if we wish, assign fixed values to certain of the a_i, by simply deleting the corresponding normal equations.)

In almost all cases of interest the normal equations are non-linear in the a_i and must be solved iteratively. *Newton's method* depends on the fact that if $\bar{a}_i$ are approximate values of the a_i, better approximations $\bar{a}_i + x_i$ are found by solving equations (18) for the first-order corrections x_i (provided the non-linearity is not too severe), this process being repeated until the corrections

$$\sum_p (F_p - S_p)F_{pi}' + \sum_j x_j \sum_p [F_{pi}'F_{pj}' + (F_p - S_p)F_{pij}''] = 0 \qquad (18)$$

become negligible. The method requires a knowledge of both first and second derivatives of F (which may not be easy to calculate) and reasonably good initial guesses $\bar{a}_i$ (otherwise the iteration may fail to converge).

Alternative methods exist for locating the minimum of ρ in the r-dimensional parameter space, *e.g.* Gauss–Newton and gradient methods, which require only first derivatives of F, and direct-search methods, which require no derivatives.[7-9] When the initial guesses are poor, these methods are more dependable than Newton's method, but the latter is nearly always faster once the iteration starts to converge.* For this reason, some of the most effective computer programs use a combination of methods.[10] Thus, it is clearly very desirable to be able to calculate the derivatives F_i', and, if possible, F_{ij}'', which means calculating derivatives of the energy levels E_n. I have shown above how these derivatives may be accurately calculated even when the E_n can be obtained only by numerical diagonalisation of the Hamiltonian matrix.

As a simple, hypothetical example of the application of these equations to a least-squares fitting problem, suppose that the observations S are of the energy levels themselves (perhaps obtained spectroscopically) for the Hamiltonian (6). The normal equations then take the quasi-linear form (19) (assuming equal weights for the observations), which could be solved immediately (*i.e.* without iteration) were it not for the fact that the $\mathscr{E}_{nn}^{(i)}$ are determined by the eigenvector matrix $\mathscr{U}$ and hence depend indirectly on the a_i. Thus, it is still necessary to make an initial guess at the values of the parameters in order to calculate $\mathscr{U}$, and then to proceed stepwise: solving for the a_i, repeating the diagonalisation with the new a_i, using the new $\mathscr{U}$ matrix to generate a new set of normal equations, and so on until constancy is attained. However, since in most eigenvalue problems a first-order perturbation of the eigenvalues (which depend directly on the a_i) is associated with only a second-order perturbation of the eigenvectors, the implicit non-linearity of equations (19) is likely to be relatively small in its effect, so that we should expect convergence of this process (which amounts to refinement of the eigenvector matrix) to be quite rapid even with relatively poor initial guesses. This expectation is borne out in a test described below.

$$\sum_j p_{ij}a_j - q_i = 0 \qquad (19)$$

where $p_{ij} \equiv \sum_n \mathscr{E}_{nn}^{(i)}\mathscr{E}_{nn}^{(j)}$ and $q_i \equiv \sum_n \mathscr{E}_{nn}^{(i)}S_n$

(*b*) *Fitting Susceptibility Data.*—(*i*) *Isotropic case.* When F is the magnetic susceptibility, the normal equations will be severely non-linear in the a_i because of the exponential way in which the eigenvalues enter equation (1). Their solution by Newton's method then requires the second derivatives of F, and because of the presence of $\partial E_n/\partial H$ in (1), this means that third derivatives of the type $\partial^3 E_n/\partial H \partial a_i \partial a_j$ will be needed. (Fourth derivatives would be required to deal in this way with the differential susceptibility, but they are not considered here.)

The process starts with the matrix $\mathscr{H}$, calculated from (6) with trial values of the a_i, which is diagonalised numerically to obtain its eigenvalues E_n and eigenvector matrix $\mathscr{U}$. The matrices $\mathscr{E}^{(i)}$ and the derivatives E_{nij}'' and E_{nijk}''' are calculated and used to obtain the values of F and those of its first and second derivatives with respect to the fitting parameters a_i at the experimental temperatures T_p, and application of equations (18) then gives a new set of a_i which should be closer to the true solution of the normal equations. These a_i are the starting point for a second iteration which (since $\mathscr{H}$ is itself a function of the a_i) ought strictly to begin with

* The following assessments are worth quoting. On Newton's method: ' If it works at all, then it works extremely well . . . and, if a sufficiently good initial estimate of the solution can be determined, it is probably the best available method.' (C. G. Broyden, ref. 8, p. 87). On the use of analytic derivatives: ' My experience has been that this extra information (*i.e.* first derivatives) is usually extremely valuable, enabling methods to be used which give an order of magnitude improvement in the time taken to solve a problem . . . (second derivatives) will not give rise to anything like so significant an improvement . . . (but) I would recommend the evaluation of second derivatives if possible.' (R. Fletcher, ref. 8, p. 125); ' (Second derivatives) almost certainly should be used except when their computation time is extremely large compared with that of the function and gradient.' (W. Murray, ref. 8, p. 71); ' (When) analytic derivatives were introduced, the picture changed dramatically. The least-squares algorithm now ceased to experience any difficulty . . .' (J. J. McKeown, ref. 9, p. 256).

another diagonalisation, so as to obtain the E_n and $\mathscr{U}$ of the revised matrix $\mathscr{H}$. However, for the reason noted above, as long as the corrections x_i are fairly small, it is often possible to carry out several successive iterations (18) with the *same* eigenvector matrix $\mathscr{U}$, provided we update the eigenvalues E_n and those derivatives that depend on them. Thus, for a linear Hamiltonian (Case A), the first-derivative matrices $\mathscr{E}^{(i)}$, which depend only on $\mathscr{U}$ and the fixed matrices $\mathscr{H}_i$, remain unchanged until a new diagonalisation is carried out, but the eigenvalues E_n and their second and third derivatives are calculated afresh at each iteration, from equations (8)—(10). Experience with the method suggests that the optimum number of iterations before rediagonalisation is likely to be between one and ten, depending on the problem.

In Case B, the fitting proceeds similarly, except that, since the matrices $\mathscr{E}^{(i)}$ depend on the a_j for $j > i$, they must be recalculated at each iteration from equation (15).

(*ii*) *Anisotropic case.* Equation (20) represents a Hamiltonian matrix in which $a_r{}^\alpha\mathscr{H}_r{}^\alpha$ is the Zeeman term,

$$\mathscr{H}_\alpha = \sum_{i=1}^{r-1} a_i\mathscr{H}_i + a_r{}^\alpha\mathscr{H}_r{}^\alpha \qquad (20)$$

with α an index distinguishing different directions of the applied field with respect to some internal co-ordinate system fixed by one or more of the $\mathscr{H}_i$. (At least one of the fitting parameters a_i, $i < r$, is here assumed not to depend on α, otherwise the data obtained with different field directions could be fitted independently, as in the example of $Cs_2[CuCl_4]$ discussed below. Note that in a 'complete' Hamiltonian, a_r would not depend on α, but most models are parametrised to the extent of using a g factor or orbital reduction factor, which is incorporated in a_r and will not necessarily be isotropic.) In the case of axial symmetry, two calculations ($\alpha \equiv \|, \perp$) are carried out to yield $\mathscr{U}_\alpha$, $E_{n\alpha}$, and $\mathscr{E}_\alpha{}^{(i)}$, from which F_α and their derivatives are calculated. If single-crystal measurements have been made, equations (18) then include an additional summation over α, but if only powder measurements are available, the average F, F', and F'' must be calculated and inserted into (18).

(*iii*) *Parameters in the field-dependent term.* The presence of $\partial E_n/\partial H$ in equation (1) causes a slight complication when the term representing the interaction with the applied field contains, besides H, a fitting parameter such as g. How we can deal with this is indicated below in three typical cases. For simplicity it is assumed here that the zero-field Hamiltonian contains only one parameter, a_1.

1. $\mathscr{H} = a_1\mathscr{H}_1 + g\mu_B\mathbf{S}\cdot\mathbf{H}$. Letting $a_2 = gH$ and $\mathscr{H}_2 = \mu_B\mathbf{S}\cdot\mathbf{H}/H$, we have equations (21). If g is fixed, the only

$$\mathscr{H} = a_1\mathscr{H}_1 + a_2\mathscr{H}_2$$
$$\partial E_n/\partial H = (a_2/H)E_{n2}' \qquad (21)$$

least-squares parameter is a_1, but if g is to be determined, both a_1 and a_2 are least-squares parameters, and the value of g is recovered from a_2 after solution of the normal equations.

2. $\mathscr{H} = a_1\mathscr{H}_1 + \mu_B(k\mathbf{L} + 2\mathbf{S})\cdot\mathbf{H}$. If the value of k is to be determined, we use the definitions (22), giving

$$a_2 = kH,\ \mathscr{H}_2 = \mu_B\mathbf{L}\cdot\mathbf{H}/H$$
$$a_3 = H,\ \mathscr{H}_3 = 2\mu_B\mathbf{S}\cdot\mathbf{H}/H \qquad (22)$$

equations (23). The fitting parameters are a_1 and a_2, and the value of k is recovered from a_2 at the end.

$$\mathscr{H} = a_1\mathscr{H}_1 + a_2\mathscr{H}_2 + a_3\mathscr{H}_3$$
$$\partial E_n/\partial H = (a_2/H)E_{n2}' + E_{n3}' \qquad (23)$$

3. $\mathscr{H} = k\lambda\mathbf{L}\cdot\mathbf{S} + \mu_B(k\mathbf{L} + 2\mathbf{S})\cdot\mathbf{H}$. Again, k is to be determined. With the definitions (24) we have equations

$$a_1 = k,\ \mathscr{H}_1 = \lambda\mathbf{L}\cdot\mathbf{S}$$
$$a_2 = H,\ \mathscr{H}_2 = 2\mu_B\mathbf{S}\cdot\mathbf{H}/H$$
$$c_{12} = 1,\ \mathscr{H}_{12} = \mu_B\mathbf{L}\cdot\mathbf{H}/H \qquad (24)$$

(25), the fitting parameter being a_1. It is in problems

$$\mathscr{H} = a_1\mathscr{H}_1 + a_2\mathscr{H}_2 + c_{12}a_1a_2\mathscr{H}_{12}$$
$$\partial E_n/\partial H = E_{n2}' \qquad (25)$$

like this one that the equations of Case B must be used.

Examples.—Although the effectiveness of the least-squares method could be tested by the use of arbitrary matrices and a variety of functions F, it seemed more useful to study reasonably realistic models and their magnetic susceptibilities. The examples below are typical of a number of real and hypothetical problems to which it has been applied. The calculations were done with a slightly modified version of FITIT, a program for multi-parameter, non-linear, least-squares fitting by Newton's method, which includes automatic reduction of the steps x_i whenever the iteration starts to diverge.[10b] This program requires the user to supply subroutines for the functions F, F_i', and F_{ij}'', and in the present case these subroutines were based on the equations given above. For each test, the program was supplied with the parameter-free matrices $\mathscr{H}_i$, the field strength H, trial values of the parameters, and a set of real or artificial data known to fit the model well for a particular set of parameter values. In most cases the test was preceded by a rough contouring of the goodness of fit over a generous sample of parameter space, and this is recommended as a preliminary to any use of traditional least-squares methods such as this.

1. *Copper*(II) *in a distorted tetrahedral environment.* The principal susceptibilities $K_\|$ and $K_\perp$ of the $[CuCl_4]^{2-}$ ion in $Cs_2[CuCl_4]$ have been determined by Gerloch and his co-workers, who also give the Hamiltonian matrix in a ^2D basis.[11,12,*] The Hamiltonian has the form (26), with λ and k to be determined by independent fits of $K_\|$ and $K_\perp$ from 90 to 300 K. The program confirmed

$$\mathscr{H} = \mathscr{H}_{\text{cubic}} + \mathscr{H}_{\text{tetrag.}} + \lambda\mathbf{L}\cdot\mathbf{S} + (k\mathbf{L} + 2\mathbf{S})\cdot\mathbf{H} \qquad (26)$$

the general shape of the contours of fit in ref. 12, and located minima of ρ at $\lambda = -745$, $k = 0.70$ and $\lambda =$

* The matrix in ref. 11 contains a misprint: the elements $\langle\phi_2\beta|\lambda\mathbf{L}\cdot\mathbf{S}|\phi_3\alpha\rangle$ and $\langle\phi_3\alpha|\lambda\mathbf{L}\cdot\mathbf{S}|\phi_2\beta\rangle$ should be equal to $\lambda/\sqrt{2}$.

-627, $k = 0.78$ for $K_{\parallel}$, and at λ *ca.* -590 cm^{-1}, k *ca.* 0.35 for $K_{\perp}$.

2. *Two octahedrally co-ordinated nickel*(II) *ions with axial distortions and an exchange interaction.* The Hamiltonian (27) was assumed to act on a $^3A_{2g} \times {}^3A_{2g}$

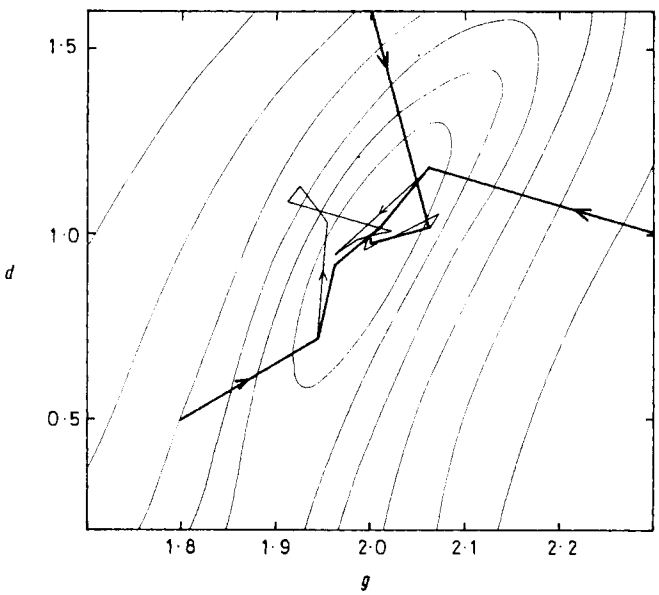
$$\mathscr{H} = D(S_{z1}^2 + S_{z2}^2 - \tfrac{4}{3}) - 2J(\mathbf{S}_1 \cdot \mathbf{S}_2) + g\mu_B \mathbf{S} \cdot \mathbf{H} \quad (27)$$

basis, giving an axial susceptibility tensor. With the reasonable assumption that the parameters D, J, and g were themselves isotropic, values of χ_x and χ_z were calculated between 1.5 and 50 K for $J/k = D/k = -2$ K and $g = 2.2$, and the average susceptibilities were then used as 'data' to be fitted. Rapid convergence to the correct parameter values was obtained from trial values deviating from the true ones by 20—30%. This is a difficult case because $\chi_{av.}$ is rather insensitive to D. A similar treatment of experimental average susceptibility data for the compound $[Ni_2(trien)_3][NO_3]_2 \cdot H_2O$ (trien = triethylenetetra-amine) between 1.5 and 60 K [13] gave a good fit with $g = 2.02$, $J/k = -1.27$ K, $D \sim 0$.

3. *Four spin-half ions coupled by equal antisymmetric exchange interactions.* This is a purely hypothetical example, in which the Hamiltonian is given by (28). If the ions lie at alternate corners of a cube, symmetry

$$\mathscr{H} = \sum_{i=1}^{3} \sum_{j=i+1}^{4} \mathbf{d}_{ij} \cdot \mathbf{S}_i \times \mathbf{S}_j + g\mu_B \mathbf{S} \cdot \mathbf{H} \quad (28)$$

requires that the vectors $\mathbf{d}_{ij}$ lie along the face diagonals connecting the vacant corners. We take the vectors to be of equal magnitude d and their signs all positive, the magnetic properties then being independent of the field direction. With the parameters $a_1 = d$, $a_2 = g\mu_B H$, the matrices $\mathscr{H}_1$ and $\mathscr{H}_2$ were calculated in $S_{12}S_{123}S$ coupling, that of $\mathscr{H}_1$ being complex. The fitting procedure was then applied to 'data' calculated for $g = 2.0$, $d/k = 1$ K, $H = 10^4$ G. Two tests were carried out. In the first, the 16 calculated energy levels were fitted directly, rapid convergence being obtained from the trial values $g = 2.5$, $d/k = 0.5$ K. In the second test, calculated susceptibilities from 1 to 100 K were used as data. The Figure shows the progress of three typical fits on a contour map of the residual function ρ. The path of convergence is seen to be relatively little affected by diagonalising only at every fifth iteration.



Contour plot of the residual function $\rho(d,g)$ of example 3. The straight lines show the progress of least-squares fits, starting from three different sets of trial parameter values, towards the solution (1.0, 2.0), with diagonalisation at each iteration (heavy lines) or every fifth iteration (lighter lines). A few attempted fits converged to a shallow subsidiary minimum at *ca.* (0, 1.9)

CONCLUDING REMARKS

The formulae given here for the derivatives of energy levels with respect to parameters of the Hamiltonian provide a general method for the calculation of properties such as the magnetic susceptibility and for the least-squares fitting of Hamiltonians to a variety of experimental data. Although such expressions are well known in the generalised perturbation theory of matrices,[14] they do not seem to have been exploited previously in problems of this kind. The restrictive assumption that all degeneracies in the energy are removed should be noted, but this assumption is often valid, and in particular it is true in experiments that involve a static magnetic field.

When the secular equation for the energy does not readily factorise (which is true in all but the simplest magnetic problems), this method offers the advantages of exact susceptibility calculation, even at temperatures of the order of $\mu_B H/k$, and of enabling fast least-squares algorithms that require derivatives of the object function to be used. In favourable cases the number of matrix diagonalisations needed in a least-squares fit can be significantly reduced, though the calculation of the higher derivatives is somewhat time consuming.

Even when the secular equation can be solved explicitly, so that algebraic differentiation of the energy expressions is possible, there may be advantages in using the matrix method, in that it allows us to replace the laborious and error-prone hand calculation and computer programming of expressions for the susceptibility and its derivatives for each individual problem by a single set of expressions which, once programmed, can be applied to a wide range of different problems, only the matrices varying from one to another.

The computations described above were performed at the Computer Centre of the University of Hull.

[0/1352 *Received*, 1*st September*, 1980

REFERENCES

[1] A. P. Ginsberg, *Inorg. Chim. Acta Rev.*, 1971, **5**, 45; E. A. Boudreaux and L. N. Mulay, 'Theory and Applications of Molecular Paramagnetism,' Wiley, New York, 1976; M. Gerloch, *Prog. Inorg. Chem.*, 1979, **26**, 1.

[2] J. H. Van Vleck, 'The Theory of Electric and Magnetic Susceptibilities,' Oxford University Press, 1932, p. 143.
[3] M. E. Fisher, *J. Math. Phys. N.Y.*, 1963, **4**, 124.
[4] Ref. 2, p. 181.
[5] (*a*) G. H. Loew, *Biochim. Biophys. Acta*, 1971, **230**, 82; V. R. Marathe and S. Mitra, *Chem. Phys. Lett.*, 1973, **19**, 140; E. König and S. Kremer, *Ber. Bunsenges. Phys. Chem.*, 1974, **78**, 268; (*b*) M. E. Lines, A. P. Ginsberg, and R. L. Martin, *Phys. Rev. Lett.*, 1972, **28**, 684; M. L. Ellzey, *Comput. Phys. Commun.*, 1975, **10**, 133; R. G. Wollmann and D. N. Hendrickson, *Inorg. Chem.*, 1977, **16**, 723.
[6] B. N. Figgis and J. Lewis, *Prog. Inorg. Chem.*, 1965, **6**, 37; F. E. Mabbs and D. J. Machin, 'Magnetism and Transition Metal Complexes,' Chapman and Hall, London, 1973, pp. 11, 130; M. Gerloch and R. F. McMeeking, *J. Chem. Soc., Dalton Trans.*, 1975, 2443.
[7] M. A. Wolfe, 'Numerical Methods for Unconstrained Optimization,' Van Nostrand Reinhold, New York, 1978.
[8] 'Numerical Methods for Unconstrained Optimization,' ed. W. Murray, Academic Press, London, 1972.
[9] 'Towards Global Optimisation,' eds. L. C. W. Dixon and G. P. Szegö, North Holland Publishers, Amsterdam, 1975.
[10] (*a*) Program MINUIT, F. James, and M. Roos, *Comput. Phys. Commun.*, 1975, **10**, 343; (*b*) Program FITIT, J. H. Head, available as program QCPE 230 from Quantum Chemistry Program Exchange, Indiana University.
[11] M. Gerloch, *J. Chem. Soc. A*, 1968, 2023.
[12] B. N. Figgis, M. Gerloch, J. Lewis, and R. C. Slade, *J. Chem. Soc. A*, 1968, 2028.
[13] A. B. Blake and W. E. Hatfield, unpublished work.
[14] P. Lancaster, 'Theory of Matrices,' Academic Press, New York, 1969, ch. 7.