

The MCGA (Multiple Cubic Gradient Approximation) Method for the Analysis of Raman Spectra

F. F. M. de Mul,* H. B. G. ten Have, C. Otto and J. Greve

Department of Applied Physics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

An easily accessible interactive method for the analysis of Raman spectra consisting of many overlapping peaks is presented. A combination of a three- or four-dimensional grid and gradient searching is applied. The method can handle spectra with up to about 50 lines, based on a broad background. Analytical and user-defined or tabulated basic functions can be included. The merits of the method are discussed with both artificial and real spectra.

INTRODUCTION

In recent years, much work has been done on the analysis and quantitative interpretation of spectra (for reviews, see Refs 1-5). Peak-fitting procedures normally apply simple functions (Gaussians, Lorentzians, Voightians) for the spectrum bands. However, in practice this turns out not to be satisfactory in all cases, since the background in the spectrum usually cannot be approximated using a simple polynomial and frequently different bands in the spectrum are coupled through the intensity, the width or the position of the bands. What one needs is a fitting procedure that is able to handle spectra, in which, apart from the simple basic functions, tabulated functions, user-defined functions or even parts of other spectra can also be used as the consisting functions. An example is found in the analysis of the complicated spectra from biological molecules, e.g. chromosomes, which contain the rich spectra of all constituents (proteins, DNA). Since no procedure in the literature seemed to meet our demands, it was decided to develop a suitable fitting procedure. This method is called 'multiple cubic gradient approximation,' since it deals with three-dimensional grid and gradient methods to search the optimum combination of parameters.

The second section contains necessary basic mathematics. In the third section we propose a straightforward, efficient and fast procedure for the analysis of complicated spectra in terms of separated bands and a well defined background: the method of 'cubic gradient approximation.' That section also deals with the computer realization of the peak-fitting procedures, together with a number of useful measures to reduce computer time and resulting errors. In the next two sections some examples are given of the analysis of spectra of varying composition, both artificial and real spectra, and of the influence of the proper definition of the background level. In the final section some conclusions concerning

the proposed method and its range of application are drawn.

BASIC THEORETICAL CONSIDERATIONS

Normally Lorentzian or Gaussian functions are used as the basic spectral functions. Measurements in condensed phases indicate that Lorentzian broadening effects are generally more frequent.⁶⁻⁸ Combinations of Lorentzians and Gaussians can also be used: linear combinations, of Lorentzians and Gaussians can also be used: linear combinations, product functions⁹⁻¹² or the Voigt convolution function:¹³⁻¹⁶

$$y(\nu) = a \int_{-\infty}^{\infty} \frac{\exp(-b\nu')^2}{c^2 + (\nu' - b)^2} d\nu' \quad (1)$$

where ν is the frequency and a , b and c are adjustable constants. Other functions, less frequently used, are inverse polynomials^{6,8,17} and combinations of basic functions.¹

The mathematical route from a noisy spectrum, consisting of broad overlapping bands to the original vibrational spectrum consists in noise filtering, deconvolution, background subtraction and peak finding and fitting:

1. *Noise filtering.* Apart from the photon noise in the signal and the background,¹⁸ the noise is produced mainly by the detector and subsequent electronics. A convenient treatment is the Savitsky-Golay multi-point averaging method.^{19,20}
2. *Deconvolution of the instrumental resolution.* This can be based on several methods: using Fourier transforms,²¹⁻²⁵ orthogonal polynomials²⁶ or probability distributions used for some classes of ill-posed problems, i.e. problems having no unambiguous solution. Recent versions of these methods are the maximum entropy method (MEM) and the Bayesian deconvolution.²⁷⁻³⁰ The latter method of deconvolution, applied in this work, will be described briefly in the subsection Deconvolution.

* Author to whom correspondence should be addressed.

3. *Background subtraction and Rayleigh-line correction.* A generally used form for the baseline is given by¹

$$y = \sum_i (a_i x^i) + b \exp(cx) \quad (2)$$

where a_i , b and c are adjustable constants. Pearson³¹ has developed a method to obtain a satisfactory polynomial initial approximation of the background of the spectrum. With this method any smoothly varying artefact can be removed, provided that the artefact is noticeably broader than the broadest cluster of overlapping spectral bands. This method has been adopted in this work. For some other studies concerning this point, see the papers by Pitha and Jones (first order, and no exponential term)¹⁰ and Gans and Gill (second order, exponential term).² Wartell *et al.*³² use polynomials up to the seventh order.

4. *Peak-search.* The centre positions of the bands can be approximated using differentiation of the spectrum. Most frequently used are first and second derivatives. The detectability limits with these derivatives have been examined by Vandeginste and De Galan.³³ Some authors even use fourth and higher order derivatives.³⁴ With each successive derivation the relative noise level of the resulting spectrum doubles,¹⁷ so it is not always justified to obtain information from derivatives of order >2 . It should be noted that a bad estimate of the background (for instance an estimation with too steep slopes) may result in a severe displacement of the zero points in the derivative spectrum of the measured spectrum and therefore in the positions of the original peaks in the spectrum also. We concluded from many simulations that an inspection by the user is still necessary between the peak-search procedures and the peak-fitting process (see Ref. 4).
5. *Peak-fitting procedures.* A number of methods have been developed for this problem (for reviews, see Refs 4, 5 and 10). Normally these methods are based on minimization of

$$\chi^2 = \sum_{i=1}^N [y_{\text{obs},i} - y_{\text{calc}}(x_i)]^2 / \sigma_{\text{obs},i}^2 \quad (3)$$

where N is the number of points in the spectrum, $y_{\text{obs},i}$ is the observed spectrum at position x_i and y_{calc} the calculated spectrum.³⁵ In photon counting experiments the standard deviation $\sigma_{\text{obs},i}$ is conveniently approximated using Poisson statistics with

$$\sigma_{\text{obs},i} = \sqrt{y_{\text{obs},i}} \quad (4)$$

The function $y_{\text{calc}}(x)$ can be expressed as

$$y_{\text{calc}}(x_i) = \sum_{j=1}^n y_{\text{band},j}(x_i) + y_{\text{backgr}}(x_i) \quad (5)$$

where j runs over the number n of bands assumed to be part of the spectrum. The functions $y_{\text{band},j}(x_i)$ and $y_{\text{backgr}}(x_i)$ denote basic spectral functions and the function fitting the background, respectively.

For the derivation of the desired parameter values by minimization of χ^2 , Bevington³⁵ distinguished two categories of methods: searching-parameter space methods (grid search or gradient search) and approximate analytical methods (parabolic expansion or χ^2 linearization). The first method takes a walk through

parameter space searching for the χ^2 minimum by scanning over the parameters in a directed or a non-directed (randomized) way. The analytical methods try to linearize the χ^2 equation by linearizing the fit functions. Gans [5] gave a comprehensive review of the applicability of numerical methods for different situations. Which method is to be preferred for spectral analysis has been examined by other workers.^{10,35,36} From arguments of accessibility and flexibility we choose a combined grid and gradient search.

MULTIPLE CUBIC GRADIENT APPROXIMATION

The approximation procedure

For the procedure of peak fitting the following assumptions have been made:

1. The centre positions of the bands in the spectrum are approximately known.
2. The spectrum consists of:
 - (a) Gaussian or Lorentzian functions for the peaks, or user-defined or tabulated functions for overlapping peaks or combination bands;
 - (b) a Gaussian or a Lorentzian function centred at zero wavenumbers for the Rayleigh line;
 - (c) a Gaussian or a Lorentzian function for the fluorescence, if present;
 - (d) a polynomial up to the fifth degree for the background (six parameters).

Based on these constraints we want a fitting procedure that:

- (a) will offer us the opportunity to interpret the fitted spectrum, the fit and the difference spectrum (= measured spectrum - calculated spectrum) during all stages of the fitting process, and the possibility of interfering by inserting or deleting bands or peaks, or altering the value of individual parameters;
- (b) is able to fit complete spectra, say up to 50 (partly overlapping) bands based upon a broad background;
- (c) will run on personal computers in reasonable times.

The method chosen is a combination of grid and gradient search, combined with an analytical method to resolve the background. For every single band j a set of k_j parameters $P_{j,i}$ is defined. In every iteration step each of these parameters is optimized by the cubic gradient search method. For every parameter an individual step size δP_i is kept in memory (for convenience we shall omit the subscript j). The best fit is calculated from the least-squares results of the fit when applying all possible parameter permutations that can be performed with the k parameters, when kept constant or when enlarged or decreased by the corresponding step size ($P_i - \delta P_i$, P_i , $P_i + \delta P_i$; see Fig. 1). This method gives 3^k possible parameter combinations (e.g. position, intensity and width). Hence for Lorentzian or Gaussian functions $3^3 = 27$ parameter combinations are possible for every band. For tabulated functions mostly the intensity and the position act as parameters only. The procedures can

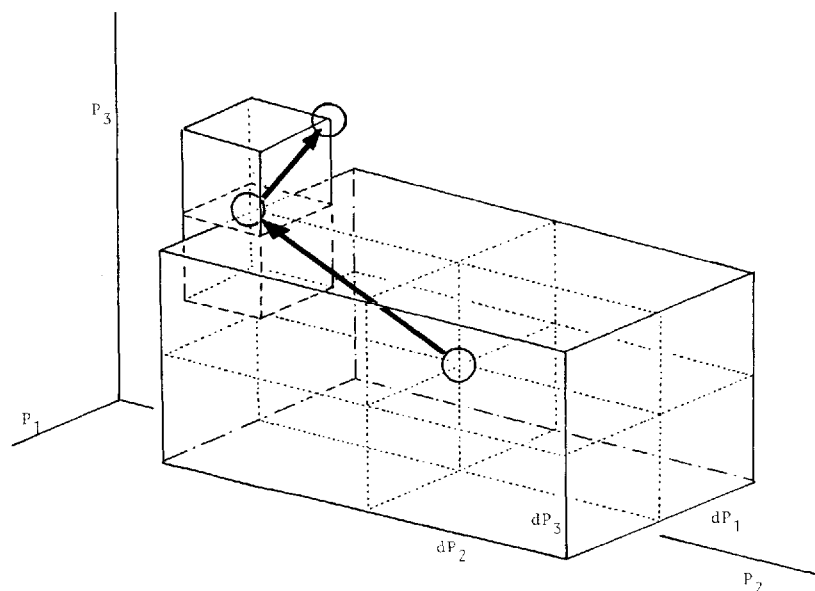


Figure 1. The cubic gradient approximation for one peak (three variables, position, amplitude and width; dP = actual step size).

easily be expanded for four-parameter functions (e.g. Voigt functions).

The reason for choosing this method of iteration, although computer time consuming,^{15,37} is its relatively simple way of handling the data, the easy way of inspection of results between cycles, the possibility of avoiding being trapped in local minima and the opportunity to intervene in the approximation process between two successive cycles. In addition, modest amounts of computer memory are needed, thus making the method available for desk-top minicomputers, commonly applied in spectroscopy. With such a computer the time needed for fitting a 40-peak spectrum, after insertion of the proper parameters, is only a few minutes. The χ^2 values that can be obtained with this method are comparable to those of other, more analytical techniques.^{4,5,10}

The background is fitted by a polynomial function.³¹ Points in the digitized spectrum $y(x_i)$ (x_i = spectrum point i) are considered to belong to the background if situated within $\pm v\sigma_b$ of the x -axis. Here v is a positive constant and σ_b is the 'baseline standard deviation':

$$\sigma_b^2 = \frac{\sum_i [y^2(v_i)h(u\sigma_b - |y(v_i)|)]}{1 + \sum_i h(u\sigma_b - |y(v_i)|)} \quad (6)$$

where v_i runs over the wavenumbers, u is a positive constant and $h(z)$ is the Heaviside function: $h(z) = 1$ if $z > 0$ and $h(z) = 0$ if $z < 0$. σ_b will converge to a well defined value if the right constants are chosen. For spectra similar to Raman spectra, Pearson³¹ found as a good choice $u = 4$ and $2 \leq v \leq 3$. Sharp-peaked signals which extend beyond the background noise level will be rejected. A smooth function (in our case a polynomial) is fitted to the background noise band ($-v\sigma_b < y(v_i) < +v\sigma_b$) and subtracted from the input spectrum. The resulting spectrum can be used for the next fit cycle.

Including functions for the Rayleigh wing (e.g. a Lorentzian centred at $v = 0$) and for the fluorescence (e.g. a broad Lorentzian centred at $v = v_{f1}$), we obtain a calcu-

lated spectrum consisting of $m + 2$ bands and a background function (a polynomial function of degree l):

$$y_{\text{calc}} = \sum_{j=1}^m y_{\text{band}j} + y_{\text{Rayleigh}} + y_{\text{fluor}} + y_{\text{backg}} \quad (7)$$

In the actual calculations we limited $m + 2$ to 50 and l to 5. In every iteration step the $(m + 2 + 1)$ sets of variables are optimized once and independently of each other by an adapted gradient search method. After each optimization of the parameters of a peak function, the corresponding step sizes are recalculated. In case the new optimum parameter value equals the value of the previous grid cycle, the step size is decreased with a predefined factor and stored in memory for use in the next grid cycle.

During the fitting procedure only the difference spectrum (= observed spectrum - calculated spectrum), the observed spectrum and the parameters of the components of the calculated spectrum have to be kept in memory. In case the procedure is started with a fit of the background, the difference spectrum is initially given by

$$y_{\text{rest}}^{0,1,k} = y_{\text{obs}} - y_{\text{calc}}^{0,1,k} \quad (8)$$

Here the first superscript indicates the number of previous fit cycles performed completely (i.e. for all constituting functions). The second is 1 or 0 depending on whether the background was optimized during or after the actual fit cycle, and the third superscript k ($k = 1-n$; $n = m + 2$) indicates the number of peaks optimized already during the actual fit.

Before the parameters of a peak can be optimized, the 'old' peak function itself has to be calculated and added to the difference spectrum. After the peak parameters have been optimized, the new peak function is calculated and subtracted from the difference spectrum. For the p th peak, y_p , in the q th cycle this results in

$$y_{\text{rest}}^{q-1,0,p} = y_{\text{obs}} - y_{\text{calc}}^{q-1,1,n} + (y_p^{q-1} - y_p^q) \quad (9)$$

where $y_{\text{calc}}^{q-1,1,n}$ is given by Eqn (7) after $q - 1$ cycles have been completed.

After optimization of all n peaks, the difference spectrum will become

$$y_{\text{rest}}^{q-1, 0, n} = y_{\text{obs}} - y_{\text{calc}}^{q-1, 1, n} + \sum_{p=1}^n (y_p^{q-1} - y_p^q) \quad (10)$$

Subsequently a background optimization is performed, which results in

$$y_{\text{rest}}^{q-1, 1, n} = y_{\text{obs}} - y_{\text{calc}}^{q-1, 1, n} + \sum_{p=1}^n (y_p^{q-1} - y_p^q) + (y_{\text{backg}}^{q-1} - y_{\text{backg}}^q) \quad (11)$$

This can be written as

$$y_{\text{rest}}^{q-1, 1, n} = y_{\text{rest}}^{q, 0, 0} = y_{\text{obs}} - y_{\text{calc}}^{q, 1, n} \quad (12)$$

by which the new difference spectrum can be used for the next iteration step. The advantage of this method is that, once the first y_{calc} has been obtained, after each step in each cycle the difference spectrum is calculated by adding the 'old' function and subtracting the new one (for all constituent functions):

$$y_{\text{rest}}^{q, 1, n} = y_{\text{rest}}^{q-1, 1, n} + \sum_{p=1}^n (y_p^q - y_p^{q-1}) + (y_{\text{backg}}^q - y_{\text{backg}}^{q-1}) \quad (13)$$

This procedure is repeated until χ^2 is reduced to a satisfactory value. It should be noted that the procedure is done peak by peak. After each peak fit in each cycle the difference spectrum can be inspected and the parameters can be changed, if necessary. If the difference spectrum had been calculated once in each cycle after the optimization of all parameter values of all peaks, problems of divergence could easily occur and as a result the difference spectrum would show strong oscillations.

MODIFICATIONS OF THE APPROXIMATION PROCEDURE

The procedure outlined above has to be modified for several reasons:

1. The possibility of becoming stuck in a non-absolute or local minimum is the major problem with gradient search methods. To overcome this problem the step sizes have to be recalculated, once a (local) minimum has been reached, to larger values, corresponding to those in a previous cycle. Two possible solutions for the recalculation procedure were tested. A first possibility is to use the step sizes that were used in the first fit cycle as the new step sizes. Secondly, it is possible to recalculate the step sizes to the values that were used by the fitting procedure just before the previous recalculation took place. Figure 2 shows the results of a fit with and without the help of the two recalculation procedures. It shows that a large gain can be made if a recalculation procedure is used. Figure 2 shows that the second procedure (curves III) gives the best results. This is due to the fact that if the first procedure is used, the step sizes are too enlarged. This is illustrated by the small step-like plateaux in the curves II.

As a result of these tests, the step sizes are recalculated in the grid cycle every time the computer

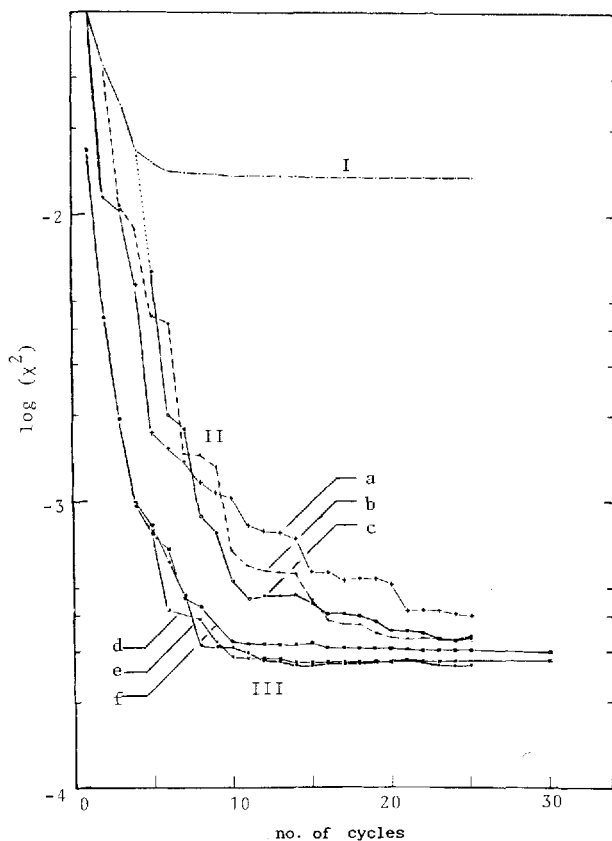


Figure 2. Effect of intermediate step size recalculation: (I) fit without recalculation; (II) fit with recalculation to initial values (a, b, c); (III) fit with recalculation to previous values (d, e, f); the recalculation is performed if the actual χ^2 is reduced by a factor of 0.75 (a, d), 0.50 (b, e) or 0.25 (c, f).

program is able to decrease χ^2 to 0.75 of the value it had during the previous recalculation (this value was found as the optimum for Lorentzian lines; in fact values between 0.25 and about 0.8 will work).

2. In order to prevent values for the peak-function parameters which are physically unrealistic, some constraints are incorporated in the program. No negative parameter values are allowed.
3. For efficiency reasons, the optimization of the parameter values is limited to a $\pm 6\sigma$ interval around the previous estimate of each peak-centre position.
4. Because most information of a band in the spectrum is concentrated in the measured points at the expected centre of the band, a modified χ^2 criterion is used, in which the difference spectrum is multiplied by a window function peaking at the calculated peak centre position. The form of the window function is given by

$$y(v) = [1 + |v_{0j} - v|/v_{0j}]^{-1} \quad (14)$$

where v_{0j} is the estimated centre position of band j . After this multiplication, the fit criterion value is calculated in the same way as normally χ^2 is calculated.

5. The first difference spectrum and the most recent difference spectrum are kept in memory and can be displayed on a monitor. Hence it is possible to inspect the changes in the difference spectrum while it reduces towards the noise spectrum, which ideally should be left after the fit is completed. Every single point of the difference spectrum can be followed

during the fit. This makes it possible to see in what regions of the spectrum the fit is troublesome, possibly owing to a bad or wrong first estimate (for instance, too few bands).

6. The order in which the different bands are fitted is important. The bands with the highest amplitudes are fitted first, because for these bands the most reliable information is available regarding position, width and amplitude. A small peak next to a large neighbouring peak in an overlapping band will tend to fit a part of its large neighbour more readily than the large neighbour will tend to do with the small peak. If the larger one is fitted first in each iteration cycle, the risk of the small peak being fitted at the expense of intensity belonging to the large neighbour is diminished. Especially when dealing with complicated spectra with many lines with often very different intensities, it turns out to be worthwhile to keep the positions of the peaks fixed at the values initially inserted, at least for a number of cycles, until a satisfactory fit with the amplitudes and the line widths as the only varying parameters is obtained.
7. Generally what is available for the parameters of the bands are approximate values for the centre position only, so the full width at half maximum intensity (FWHM) and the amplitude have to be estimated. This estimation can be done by the user (subjective) or by a programmed algorithm (the drawback is that errors in unexpected or unpredictable situations can occur). Both kinds of input are made possible in the program. To start the fit close to an expected χ^2 minimum, the peak-function parameters FWHM and amplitude can be estimated from the measured spectrum. Therefore, at first the background is estimated by the algorithm described [Eqn (6)]. After subtraction of the background, the expected peak positions are imported by the program. As a first approximation of the amplitude the values measured at the given positions, corrected for background, are taken. To obtain the FWHM values of the peaks constituting a multi-peaked band the program searches for the width at a suitably chosen percentage $X\%$ of the amplitude, i.e. the width of the interval (not necessarily symmetrical) around the estimated centre of the band, comprising points with measured intensities of more than $X\%$ of the estimated band amplitude. The first FWHM approximation is the width of the interval found, divided by the number of peak-centre positions (imported in the program) expected within the interval. The individual amplitudes are recalculated, based on the approximations found for the width of the peak functions and the measured amplitudes at the estimated peak-centre positions. Empirically and with help of a triangular model we found that $X > 80\%$ gives good results (see the next section).

Initial approximation of peak widths

The fitting procedure starts with the input of the number of peaks, estimated 'with a spectroscopists view' or using spectrum derivatives together with a reasonable initial guess (to be discussed later) of the spectral position of each peak. The computer program has to

calculate the first estimations of the peak amplitude and of the peak width.

In order to investigate the applicability of the procedure already discussed in the previous section (point 7) for broad overlapping bands consisting of two or more peaks, a number of test spectra were analysed. The test spectra consisted of peaks of various amplitude, width and spectral separation. For convenience these peaks were chosen to be triangular instead of Lorentzian, since in this case the width of a single Lorentzian is satisfactorily approximated using a triangular function, provided that the approximation is based on the upper part of the Lorentzian (relative amplitude $> 80\%$). This is illustrated in Fig. 3(a) and (b), where for a Lorentzian and a triangular peak the width, approximated using the width at a level of $X\%$ of the intensity in the presumed peak position, is compared with the real width. The results for various combinations of triangular peaks are displayed in Fig. 3(c) and (d). With these figures it was assumed that $\Gamma_1, \Gamma_2 < |\mu_2 - \mu_1| < \Gamma_1 + \Gamma_2$, (Γ_i and μ_i are the width and position, respectively, of peak i ; thus allowing overlap). Then it can be shown that the width of each peak equally can be approximated by Γ' , with

$$2\Gamma' = (\Gamma_2 + \mu_2)(1 - X) + Xy - \Gamma_1(X' - 1) - \frac{\Gamma_1}{\Gamma_2} X'(\mu_2 - y) - \mu_1, \text{ for } y > \mu_2$$

where $X' = X A_2 / A_1$ ($A_i =$ amplitude of peak i) and y is the running variable, and with similar expressions for other y values. At the start of the fitting procedure y is chosen as the estimated position of the supposed highest peak in the band. It is seen that with this pro-

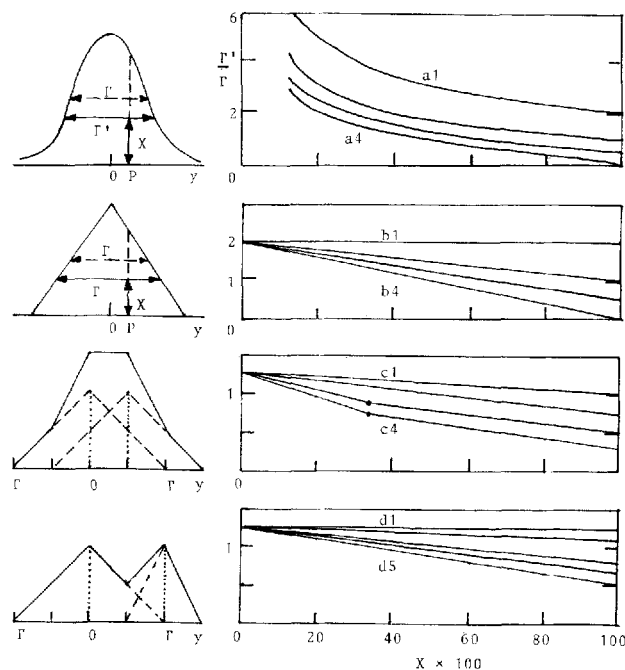


Figure 3. Approximation of the widths of two overlapping peaks using triangular line shapes. P denotes the estimated peak position; Γ is the real FWHM and Γ' the estimated FWHM, using a fraction X of the intensity at P . Parameter: y_p . Values of y_p : (a) Lorentzian line: a1, Γ ; a2, $\Gamma/2$; a3, $\Gamma/4$; a4, 0; (b) triangular peak: b1, Γ ; b2, $\Gamma/2$; b3, $\Gamma/4$; b4, 0; (c) and (d) two overlapping triangular peaks: c1, $5\Gamma/4$; c2, Γ and $-\Gamma/2$; c3, $3\Gamma/4$; c4, $\Gamma/2$ and 0; d1, 1.5Γ and $-\Gamma$; d2, 1.4Γ and -0.8Γ ; d3, 1.2Γ and -0.4Γ ; d4, 1.1Γ and -0.2Γ ; d5, Γ and 0.

cedure a reasonable guess of the widths of the peaks is obtained. As will be discussed later, in the case of overlapping bands the central positions should be initially guessed within 10% of the width, and therefore in the region where $-\Gamma/10 \lesssim (x \text{ or } y) \lesssim +\Gamma/10$, where Γ is the FWHM and x and y are the running spectral variables of both peaks. The figure shows the ratio of the estimated width Γ' and the real width Γ vs. the relative amplitude X (as a percentage of the band amplitude) that the data points should have at least in order to be contained in the calculation. From the figures the optimum value of X may be derived.

Error calculation for the resulting parameters

The commonly used methods for the estimation of the errors in the parameter values resulting from the fitting procedure are a variation of the initial parameter values (see, e.g., Ref. 32) and a variation of the final parameter values by varying the resulting step sizes. The first method can give rise to large, unrealistic, values of the relative errors in the parameters,³² up to 20%. Some problems with this method are a lack of convergence and the possibility of 'excursions' of peaks to very broad bands, adding to the background.

The second method, adopted in this work, consists of subsequently varying the parameters P_i over a fairly large region. This is achieved by adding $\pm 2^n \epsilon_{if}$ to P_{if} , where $n = 1, 2, \dots$, P_{if} is the final value of the parameter P_i after completion of the fitting procedure and ϵ_{if} is the corresponding final step size. By repetitive scanning of the P_i axis in this way, keeping the other parameters constant, the value of P_i with minimum χ^2 (P_{im}) can be found [Fig. 4(a)]. Then the estimation of the error in P_i is given by

$$\Delta P_i = |P_{im} - P_{if}| \quad (15)$$

This method contains some inherent traps. Two of those are illustrated in Fig. 4: (a) a mismatch of the final step size to the actual χ^2 curve and (b) the disturbing effect of non-independent parameters as with overlapping bands. Therefore, this method will work only close to the absolute minimum value of the χ^2 function.

Deconvolution

Of the methods available for the deconvolution of spectra we adopted³⁸ the method based on the postulate of Bayes, as described by Kennett and co-

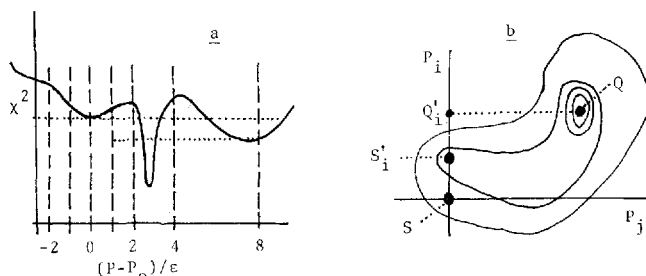


Figure 4. Erroneous calculation of the error in the variable P , caused by: (a) a mismatch of the scan procedure; (b) mutual dependence of parameters. S is the optimum parameter combinations and S' is the minimum found with (a). The error in P_i is $|Q'_i - S_i|$ rather than $|S'_i - S_i|$.

workers.²⁷⁻³⁰ For convenience we shall give here a short account of this method. The essence of the postulate is the statement

$$P(A|B) = P(B|A) \frac{P(A)}{P(B)} \quad (16)$$

in which $P(A)$ and $P(A|B)$ denote *a priori* and *a posteriori* probabilities, respectively:

$$P(B) = \sum_a P(B|a)P(a) \quad (17)$$

Writing \mathbf{M} ($M_i, i = 1, \dots, N$, $N =$ number of points) for the measured spectrum, \mathbf{O} for the original spectrum and \mathbf{C} for the convolution matrix, one can write

$$P(M_j) = \sum_i P(M_j|O_i) \cdot P(O_i)$$

and

$$P(O_i|M_j) = \frac{P(M_j|O_i) \cdot P(O_i)}{\sum_k P(M_j|O_k) \cdot P(O_k)} \quad (18)$$

where the probabilities $P(a_i)$ (a is an arbitrary spectrum) are given by

$$P(a_i) = a_i / \sum_j a_j$$

and since $\mathbf{M} = \mathbf{C} \cdot \mathbf{O}$, it is seen that $P(M_i|O_j) = C_{ij}$ and $P(O_i) = \sum_j P(O_i|M_j) \cdot P(M_j)$. Therefore,

$$P(O_i) = \sum_j \frac{P(M_j|O_i)P(O_i)P(M_j)}{\sum_k P(M_j|O_k)P(O_k)} \quad (19)$$

The terms $P(O_k)$ are not known on the beforehand and have to be estimated. A suitable procedure is an iterative relationship:

$$P_{m+1}(O_i) = \sum_j \frac{P(M_j|O_i)P_m(O_i)P(M_j)}{\sum_k P(M_j|O_k)P_m(O_k)} \quad (20)$$

m being the number of iterations. Writing out the probabilities, one finds

$$O_i^{(m+1)} = O_i^{(m)} \sum_j \frac{C_{ji} M_j}{\sum_k C_{jk} O_k^{(m)}} \quad (21)$$

With this relationship the original spectrum may be approximated iteratively. However, a problem arises upon implementation of this algorithm: a spectrum of N points requires about N^4 multiplications per iteration cycle. This amount is reduced by the rather diagonal character of the \mathbf{C} -matrix, since mostly the broadening function extends over a limited number (say $-M, \dots, +M$) of spectral points only. Then the summations over j and k extend over $i - M, \dots, i + M$ and $j - M, \dots, j + M$, respectively. The number of multiplications is now reduced to $N(4M + 1)$. Some corrections at the borders of the spectrum are necessary.

Over- or super-deconvolution. A problem often encountered when applying Eqn (21) is caused by the lack of a proper stop mechanism: with test spectra it was seen that the iterative process is indeed effective in reproducing the original spectra, but the process will not stop by itself when the original spectrum is reproduced.³⁸ A

further sharpening of peaks is obtained. This effect mainly originates from the fact that the deconvolution process will handle spectral regions with a low value for the slope more effectively (on a time scale basis) than steep regions. The deconvolution of the peak region of a spectral band is completed only after the completion of the deconvolution of the wings. When using a uniform deconvolution procedure for all points in the spectrum, the wings are then easily over-deconvoluted.

However, when in Eqn (21) $O_i^{(m)}$ is replaced by $O_i^{(0)}$, it turns out that with the test spectra investigated the iteration procedure leads to a convergent result, lying between **M** and **O** (dependent on the shape of the spectrum). It may therefore be expected that a suitable mechanism of damping the convergence and avoiding superresolution may consist of replacing $O_i^{(m)}$ of Eqn (21) by

$$(1 - d)O_i^{(m)} + dO_i^{(0)} \quad \text{with } 0 < d < 1 \quad (22)$$

where d is dependent on the actual degree of convergence obtained during the process. It can be shown that a suitable damping is obtained with

$$\begin{aligned} d &= 0, & \text{for } |1 - O_i^{(m+1)}/O_i^{(m)}| &\geq \eta \\ d &= d_0, & \text{for } |1 - O_i^{(m+1)}/O_i^{(m)}| &< \eta \end{aligned} \quad (23)$$

respectively, where d_0 is a properly chosen damping factor and η the overall accuracy to be pursued. The deconvolution process is stopped after the iteration cycle in which the number of points in the spectrum, upon which damping is applied, for the first time exceeds a limiting value of $d_f N$, where d_f is an error limit given by the minimum value of $\frac{1}{2}$ and d_0/η and N is the number of points in the spectrum.

Typical values are $\eta = 10^{-2}$ and $d_0 = 0.1$, so $d_f = 0.1$. In Fig. 5 some results of the deconvolution procedure, including the effect of over-deconvolution, are shown.

Realization

The computer program is written in Pascal, in modular form, for MS-DOS. In addition to the proper fitting procedures, a number of procedures concerning the handling of data files (i.e. intensity and spectral region corrections, procedures for calculations and plot and print routines) have been added.

TESTS OF THE PROCEDURES WITH ARTIFICIAL SPECTRA

In order to investigate the quality and the shortcomings of the method, a number of tests were performed. This section deals with the treatment of artificial spectra and 'real' spectra are discussed in the next section.

A comparison with literature data is difficult since the various studies differed in methods, initial parameter values, maximum likelihood criteria or noise contribution. Maddams⁴ and Gans⁵ gave critical reviews and Pitha and Jones¹⁰ compared different methods concerning the necessary computing time.

The artificial spectra used consisted of one, two or three Lorentzian lines, smooth or noisy. The noise was Poisson noise, calculated for a default mean value of

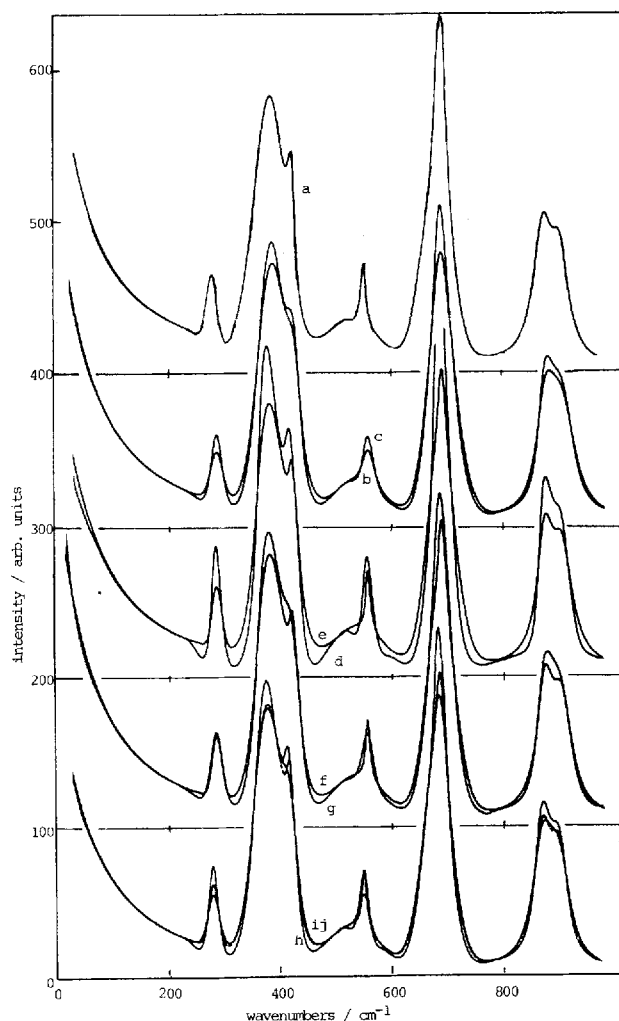


Figure 5. Over-deconvolution. Example of the damping procedure. The wavenumber axis can be re-scaled. (a) The original spectrum (artificial: Lorentzian lines at 0 cm^{-1} and at indicated positions); (b) convolved with a Lorentzian function (FWHM = 20 cm^{-1}); (c) deconvolution without damping, after 10 cycles (incomplete); (d) as in (c), after 20 cycles (over-deconvolution); (e) = (a); (f) deconvolution with damping with 100% original spectrum [$d = 1$ in Eqn (22)]: automatic stop (χ^2 criterion) after 10 cycles; (g) = (a); (h) deconvolution with mixed damping [Eqn (22): $d = 0.1$]: automatic stop after 40 cycles, showing a close resemblance to the original spectrum (noteable differences in the top intensity of the highest lines at about 400 and 700 cm^{-1} : -6 and -10 a.u. respectively; differences in the peak positions and widths <2%); (i), (j) forced continuation of the deconvolution process with a damping as in (h), showing increasing loss of sharpness in the peaks.

10^4 (with a standard deviation of 10^2), using the Stirling approximation for $n!$ to reduce the computing time. Each point y_i in the smooth spectrum is corrected for the noise contribution, calculated by multiplying the noise level σ_i ($\sigma_i^2 = y_i$) with the proper probability factor resulting from a random choice in the Poisson spectrum.

With the artificial spectra consisting of two or more peaks, the effects of a variation of a number of parameters were studied, such as the relative positions, amplitudes and widths of the peaks and the error estimations of those parameters as a function of the number of cycles, as compared with the actual values of those parameters chosen at the beginning of the fitting pro-

cedure, which generally differ from the exact values. Also the error estimation in the parameters will differ from the exact losses.

It is convenient to use the following notation:

\tilde{X}_i = exact parameter value of peak i ($X = P, A$ or F for the position, the amplitude or the full width at half maximum, respectively);

X_i = approximated parameter value after the fitting procedure;

X_i' = approximated parameter value at the start of the fitting procedure;

δX_i = the error in X_i , estimated after the fitting procedure;

ΔX_i = the exact error in X_i : $\Delta X_i = X_i - \tilde{X}_i$.

The distance between two peaks has been normalized to D , with $D = dF_j/F_i$, where d is the real distance and $F_j/F_i \geq 1$.

The errors ΔP_i and ΔF_i were normalized to F_i and the error ΔA_i was normalized to A_i . The peak positions at the beginning of the fit were treated in a relative way by normalization of the distance between the exact position and the first estimate to the width of the peak, i.e. $dP_i = (P_i - P_i')/F_i$.

Of the large number of spectra only a few examples will be discussed here. In Fig. 6 a survey is given of the results of the fit of two overlapping peaks (the parameters χ^2 , ΔP_2 , ΔA_2 , ΔF_2) as a function of the initial error in the position of one of the peaks (dP_2), with the amplitude ratio A_2/A_1 and the initial position error of the other peak (dP_1) as the parameters. The broken lines in Fig. 6 are merely drawn to guide the eye. Smooth spectra with no noise were used here. From Fig. 6 it is seen that an increase in dP_2 generally results in a worse fit. However, a negative value for dP_2 improves the fit, provided that $|dP_1| < |dP_2|$. This can be shown to be caused by the better first estimate of the combined spectrum: the latter spectrum resembles the exact spectrum more closely. This is also the reason

why the plots in the asymmetric case ($A_2/A_1 < 1$) seem to be shifted to higher dP_2 values as compared with the plots in the symmetric case ($A_2/A_1 = 1$), revealing minima at about $dP_2 \approx +10\%$.

From these and similar plots it can be concluded that in order to obtain errors in the parameters of both peaks below 5%, it is necessary that dP_1 and dP_2 are limited to $-10\% < dP_1 < 0\%$ and $0 < dP_2 < 10\%$, in which it is assumed that peak 1 is at the negative side of peak 2, and $A_2/A_1 = 1$. When A_2/A_1 is increased, the above range of dP may also be increased.

In Fig. 7 some results are shown concerning fits as a function of the initial guess for the distance between the positions of the bands. In the plots the effect of Poisson noise is also indicated: (a1)–(a4) for smooth spectra and (b1)–(b4) for noisy spectra. Comparing Figs 7(a1) and (b1), it is seen that with smooth spectra different values of the amplitude ratio only slightly influence the final result, whereas with noisy spectra a reliable result can be obtained only for large d values. With both types of spectra the best results can be expected when $D \approx 0.75$, whereas with Poisson noise only D values larger than 1 give error results $< 2\%$.

These and similar figures can be used for the estimation of accuracies in fits with real spectra. For instance, in order to obtain fit errors in the parameters of $< 15\%$ when fitting two partly overlapping peaks with equal widths, the amplitude ratio A_2/A_1 should exceed 0.75 (when $D \approx 0.5$) and 0.25 (when $D \approx 1$). When the width ratio F_2/F_1 decreases from 1.0 to 0.5, the error to be expected to increase by a factor of 4–6.

ANALYSIS OF COMPLICATED SPECTRA

Methyl alcohol–ethyl alcohol mixtures

In order to test the features of the fitting procedure outlined above in a more quantitative way, some mixtures

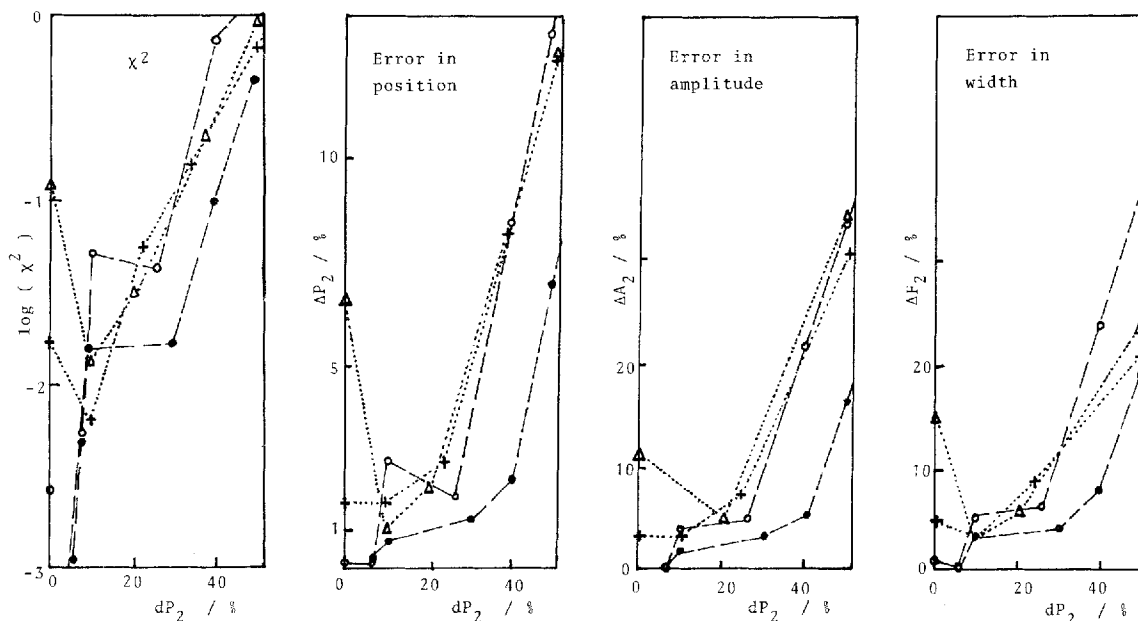


Figure 6. Fit results of test spectra with two Lorentzian lines. Shown: χ^2 value and the errors in position, amplitude and width of line 2 on variation of the initial guess error of the position. Amplitude ratio $A_2/A_1 = 1.00$ (\circ , \bullet) and 0.75 ($+$, Δ). Parameters: $dP_1 = 0$ (\circ , $+$) and -10% (\bullet , Δ); $F_2/F_1 = 1.00$; normalized band spacing $D = 0.8$; number of cycles = 15. dP , ΔP = error in initial and final parameter value, respectively.

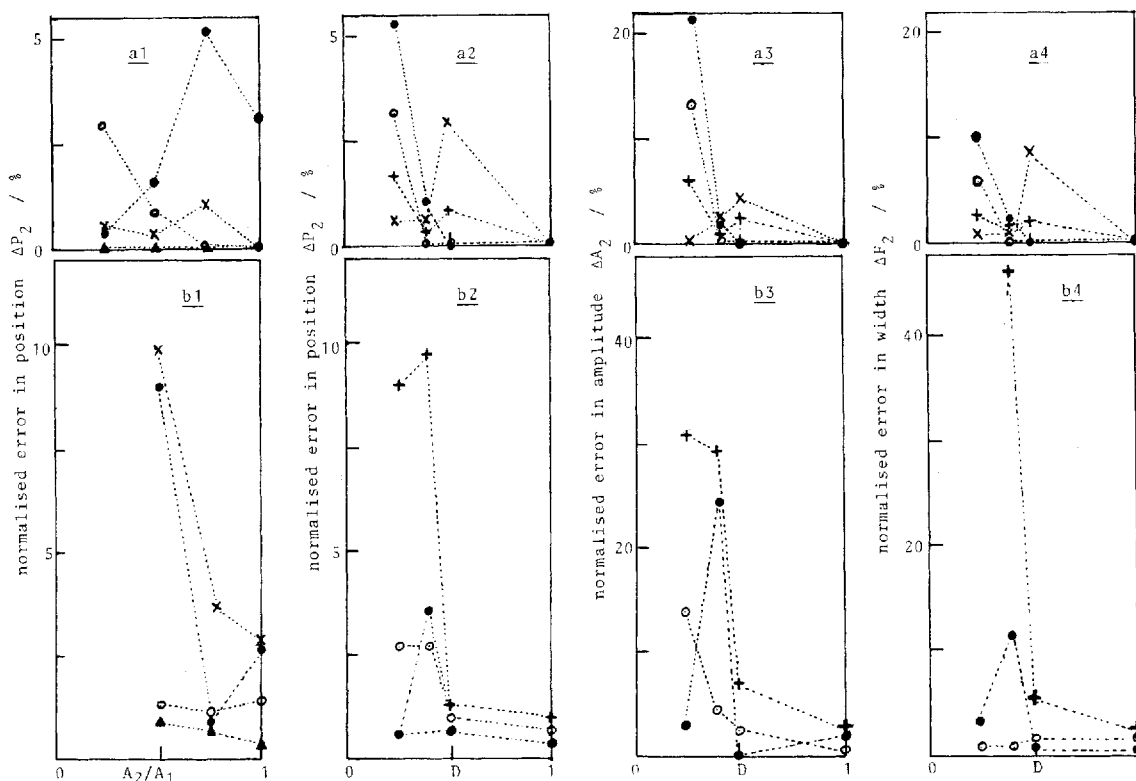


Figure 7. Fit results of test spectra with two Lorentzian lines. Shown: normalized errors in position, amplitude and width as a function of the amplitude ratio A_2/A_1 (parameter: band spacing D) and the band spacing D (parameter: amplitude ratio). (a) Smooth spectra; (b) Poisson noise included. In (a1) and (b1): $D = 0.5$ (●), 0.8 (×), 1.0 (○) and 2.0 (▲). In the other plots: $A_2/A_1 = 0.25$ (×), 0.50 (+), 0.75 (●) and 1.00 (○). Other parameters: $F_2/F_1 = 1.00$; number of cycles = 30; $dP_1 = 0\%$; $dP_2 = 10\%$.

of methyl alcohol and ethyl alcohol with various concentration ratios were measured and the spectra analysed. In Fig. 8 the spectra of the two compounds and the mixtures are shown for the CH and OH stretching region around 3000 cm^{-1} . For the calculations the spectra were calibrated using the 881 cm^{-1} line of ethyl alcohol, which is situated in the spectrum in a satisfactorily isolated way, where the contribution of methyl alcohol to the spectrum is comparatively negligible.

The methyl alcohol and ethyl alcohol lines to be expected in this spectral region are shown in the inset in Fig. 8. The fitting region extended from 2780 to 3100

cm^{-1} . It is seen from the spectra that the methyl alcohol peak in particular at 2939 cm^{-1} seems to have disappeared. In such situations it turns out in practice that a peak initially situated in the wing of a (relatively) giant neighbour will, during the fitting process, either shift to the position of this large neighbour or become very broad (thus becoming part of the background), or both, unless the position of the peak(s) are kept fixed for a limited number of cycles.

In Fig. 9 the results for the various lines are shown as a function of ethyl alcohol concentration. The lines drawn through the measured points indicate the expected relationship. It is seen that the weak (and disappearing) peak at 2939 cm^{-1} is easily reconstructed.

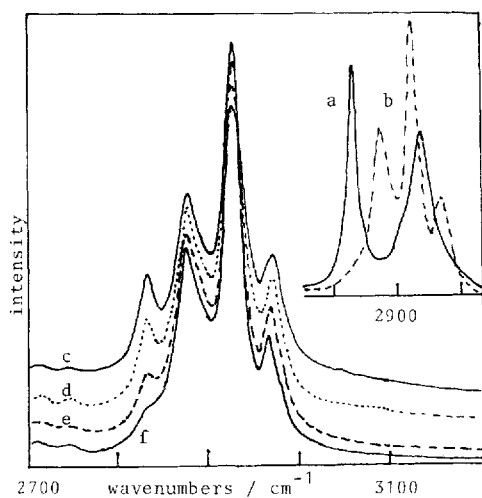


Figure 8. Methyl alcohol-ethyl alcohol spectra; CH and OH stretch region. (a)-(f): 100, 0, 75, 80, 89, 95% ethyl alcohol. Measured and fitted spectra coincident; 10 cycles.

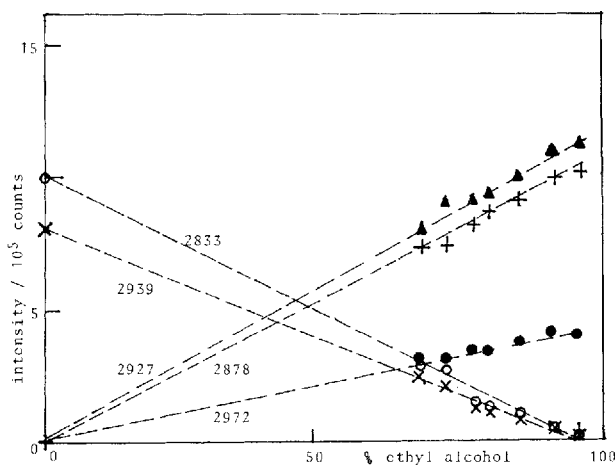


Figure 9. Peak intensities resulting from the fits corresponding to Fig. 8. Parameter: position of ethyl alcohol or methyl alcohol lines.

Human eye lenses

The spectral region mentioned above is also of interest in the investigation of the water content of human eye lenses, supposedly connected with cataract formation.^{39,40} In Fig. 10 a typical spectrum is shown, together with the fit of the NH, CH and OH stretching region. The quality of the fit is revealed in the difference spectrum. It turns out that the water band may be approximated with three Lorentzians (we shall not discuss the physical relevance of this assignment here). The background was taken as linear, on the basis of the shape of the spectral regions (not shown) at the left and right of the region shown.

Chromosomes

In Fig. 11, part of a spectrum of pelleted chromosomes is shown.^{41,42} It is seen that the measured and calculated spectra coincide. The background consisted of a polynomial with degree 4. All lines can be assigned, including the amide lines around 1600 cm^{-1} .

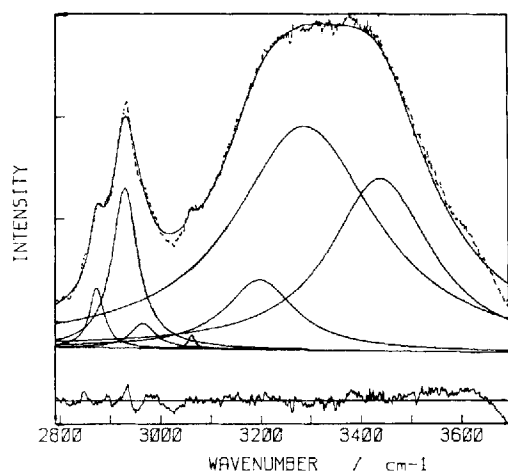


Figure 10. Micro-Raman spectrum of a human eye lens slide for cataract research; determination of the absolute water content.^{39,40} NH at 2860 cm^{-1} CH at 2950 cm^{-1} , OH at 3180–3430 cm^{-1} . Plotted: measured, calculated and difference spectra.

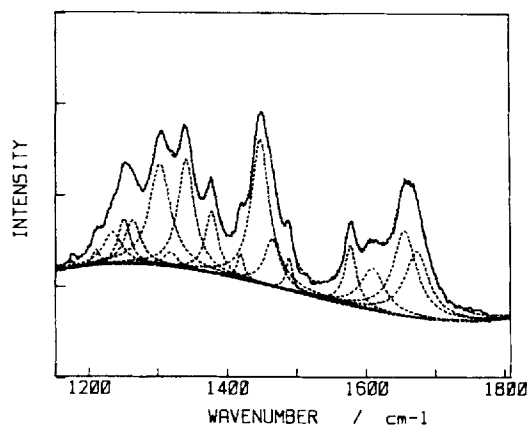


Figure 11. Micro-Raman spectrum of pelleted chromosomes in PBS.^{41,42} The measured and calculated spectra coincide. All peaks can be assigned.

AHP

Figure 12 shows spectra of AHP [ethyl hydroxyphosphate, $\text{C}_5\text{NH}_9(\text{PO}_3\text{H}_2)_2$], which are seen to consist of at least 35 peaks. The agreement between the measured and calculated spectra after five cycles is satisfactory (1 cycle = 35 subsequent peak fits + 1 background fit). The fits took about 5 min each. At this stage in the fitting procedure some small corrections have to be made: the line at $640.3 \pm 0.2 \text{ cm}^{-1}$ is slightly Gaussian; the line at $751.2 \pm 0.2 \text{ cm}^{-1}$ has two small shoulders (not yet included in the fit); and additional lines have to be inserted at about 560 and 1404 cm^{-1} . The intensity of these additional lines can be inserted from the intensity in the difference spectrum at these points.

The spectra shown in Fig. 12 were chosen specially with no side wings present in order to judge the degree of the background polynomial. We therefore tested the fitting procedure with different background degrees (0–4, Fig. 12(a)–(f)). From the results it follows that with degree 0 to 1 similar results are obtained, except for some small differences in the amplitude or the width of some individual peaks (but not the position or the integrated intensity). Background degree 2 leads to an even better result. With degree 3 the result becomes worse, with a tendency for low-intensity peaks to broaden considerably, in order to compensate for a background which apparently is too low locally (e.g. the peaks at 380 and 580 cm^{-1}). This tendency is even more manifest with degree 4 at the low-wavenumber side of the spectrum. It can therefore be concluded that additional information on the background (e.g. about the side wings of the spectrum) is of paramount importance for a successful fit.

The errors in the line positions vary from 0.2 cm^{-1} for the sharp and isolated peaks to about 3 cm^{-1} for the low and broad bands. The errors in the integrated intensity vary from 5% (sharp lines) to about 60% (low broad bands).

DISCUSSION AND CONCLUSIONS

The MCGA method for the analysis of complicated spectra has been described, and a number of applications have been given, together with some results from the analysis of test spectra. In all cases a satisfactory fit was obtained, although in some cases a continuation of the fitting process seems appropriate, especially when extra peaks of low intensity have to be included in the process. Additional knowledge of the background function is very important. The time required for the fit (about ten cycles) is a few minutes on a PDP-LSI 11/73 computer (comparable to an AT-type PC) for spectra consisting of 35 peaks.

An important advantage of this direct method is the opportunity to interfere within the process, e.g. for the insertion of extra lines, or for manipulation of the degrees of freedom. Another advantage is the opportunity to insert user-defined or tabulated functions, e.g. results from previous measurements. As a general point it may be concluded that, especially with

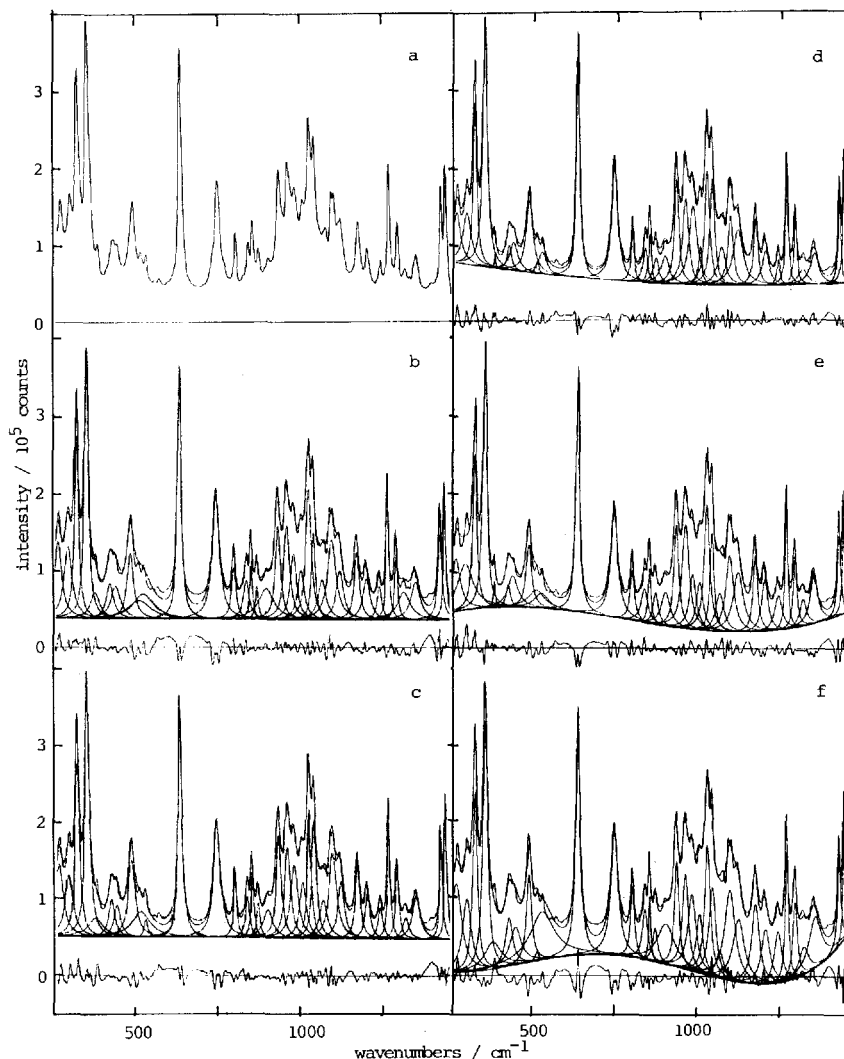


Figure 12. Micro-Raman spectrum of AHP [ethyl hydroxyphosphate, $C_5NH_9(PO_3H_2)_2$]. 800 points; 35 Lorentzian lines. (a) Measured spectrum; (b)–(f) measured and calculated spectra, the difference spectrum (= measured – calculated spectrum), the background and the constituent lines. Background: polynomial degree: (b)–(f) = 0–4, respectively. Fit results (χ^2): (b) 702; (c) 447; (d) 384; (e) 556; (f) 941 (arbitrary units).

complicated spectra, such as those for biological molecules, the opportunity to intervene in the process is necessary, in order to remove unrealistic lines or to add apparently needed new lines.

Acknowledgement

Thanks are due to J. C. de Knijff, who provided the deconvolution procedures, and who also enabled us to use the computer system in the most efficient and convenient way.

REFERENCES

1. D. C. Price, *Aust. J. Phys.* **35**, 51 (1981).
2. P. Gans and J. B. Gill, *Appl. Spectrosc.* **31**, 451 (1977).
3. A. Baruya and W. F. Maddams, *Appl. Spectrosc.* **32**, 563 (1978).
4. W. F. Maddams, *Appl. Spectrosc.* **34**, 245 (1980).
5. P. Gans, *Coord. Chem. Rev.* **19**, 99 (1976).
6. C. Baker, W. F. Maddams, J. G. Grasselli and M. A. S. Hasle, *Spectrochim. Acta, Part A*, **34**, 85 (1976).
7. F. J. Bartoli and T. A. Litovitz, *J. Chem. Phys.* **56**, 244 (1968).
8. C. Baker, P. S. Johnson and W. F. Maddams, *Spectrochim. Acta, Part A* **34**, 683 (1976).
9. S. Abramowitz and R. P. Bauman, *J. Chem. Phys.* **39**, 2757 (1963).
10. J. Pitha and R. N. Jones, *Can. J. Chem.* **44**, 3031 (1966).
11. D. Papousek and J. Pliva, *Collect. Czech. Chem. Commun.* **30**, 3007 (1965).
12. R. P. Young and R. N. Jones, *Chem. Rev.* **71**, 219 (1971).
13. P. Martin and J. Puerta, *Appl. Opt.* **20**, 259 (1981).
14. J. Puerta and P. Martin, *Appl. Opt.* **20**, 3923 (1981).
15. R. J. Noll and A. Pires, *Appl. Spectrosc.* **34**, 351 (1980).
16. A. Klim, *J. Quant. Radiat. Transfer* **26**, 537 (1981).
17. T. C. O'Haver and G. L. Green, *Anal. Chem.* **48**, 312 (1976).
18. W. Pitz and H. Kriegsmann, *Z. Phys. Chem.* **268**, 215 (1987).
19. A. Savitsky and M. J. E. Golay, *Anal. Chem.* **36**, 1627 (1964).
20. J. Steiner, Y. Termonia and J. Deltour, *Anal. Chem.* **44**, 1906 (1972).
21. W. R. Bennett, *Appl. Opt.* **17**, 3344 (1978).
22. J. Low and M. C. Jones, *Nucl. Instrum. Methods* **225**, 127 (1984).
23. R. Verma, *Nucl. Instrum. Methods* **212**, 323 (1983).
24. F. Ni and H. A. Scherag, *Anal. Chem.* **36**, 1627 (1964).
25. P. Verkerk, *Comput. Phys. Commun.* **25**, 325 (1982).
26. L. M. Delves and J. Walsh, *Numerical Solutions of Integral Equations*. Clarendon Press, Oxford (1974).
27. T. J. Kennett, W. V. Prestwick and A. Robertson, *Nucl. Instrum. Methods* **151**, 285 (1978).
28. T. J. Kennett, W. V. Prestwick and A. Robertson, *Nucl. Instrum. Methods* **151**, 293 (1978).

29. T. J. Kennett, P. M. Brewster, W. V. Prestwick and A. Robertson, *Nucl. Instrum. Methods* **153**, 125 (1978).
30. T. J. Kennett and W. V. Prestwick, *Nucl. Instrum. Methods* **203**, 317 (1982).
31. G. A. Pearson, *J. Magn. Reson.* **27**, 265 (1977).
32. R. M. Wartell, J. T. Harrel, W. Zacharias and R. D. Wells, *J. Biomol. Struct. Dyn.* **1**, 83 (1983).
33. B. G. M. Vandeginste and L. de Galan, *Anal. Chem.* **47**, 2124 (1975).
34. W. L. Butler and D. W. Hopkins, *Photochem. Photobiol.* **12**, 439, 451 (1970).
35. P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill, New York (1969).
36. B. J. Duke and T. C. Gibb, *J. Chem. Soc. A.* **68**, 1478 (1967).
37. J. Becksey, L. Berke and J. R. Callan, *J. Chem. Educ.* **45**, 728 (1968).
38. J. C. De Knijff, Internal Report, University of Twente, Department of Applied Physics, Enschede (1986).
39. A. Huizinga, A. C. C. Bot, F. F. M. de Mul, G. F. J. M. Vrensen and J. Greve, *Exp. Eye Res.* **48**, 487 (1989).
40. A. C. C. Bot, A. Huizinga, F. F. M. de Mul, G. F. J. M. Vrensen and J. Greve, *Exp. Eye Res.* **49**, 161 (1989).
41. G. J. Puppels, C. Otto, F. F. M. de Mul and J. Greve, in *Proc. International Conference on Raman Spectroscopy—XI*, edited by R. J. H. Clark and D. A. Long. Wiley, Chichester (1988).
42. G. J. Puppels, F. F. M. de Mul, C. Otto, J. Greve, M. Robert-Nicoud, D. A. Arndt-Jovin and T. M. Jovin, *Nature (London)* in press.