Generalized spectral decomposition method applied to infrared, ultraviolet, and atomic emission spectrometry. Anal. Chem. 48:1540-1546

ARTICLE *in* ANALYTICAL CHEMISTRY · SEPTEMBER 1976

Impact Factor: 5.64 · DOI: 10.1021/ac50005a031

3 AUTHORS, INCLUDING:

Harvey Gold
Dupont
30 PUBLICATIONS **435** CITATIONS
SEE PROFILE

C. Rechsteiner
CRechsteiner Consulting LLC
24 PUBLICATIONS **150** CITATIONS
SEE PROFILE

(3) "Official Methods of Analysis of the Association of Official Agricultural Chemists", 10th ed., Association of Official Agricultural Chemists, Washington, D.C., 1945, pp 354–358.
(4) "U.S. Pharmacopeia XVII", Pharmacopeia of the U.S., 1971, pp 894–895.
(5) L. Moreau and E. Vinet, *C. R. Hebd Seances Acad. Sci.*, **158**, 869 (1914).
(6) A. F. Judd, *J. Am. Pharm. Assoc.*, **2**, 961 (1913).
(7) L. V. Markova and T. S. Maksimenko, *Zh. Anal. Khim.*, **25**, 1620 (1970).
(8) D. Liederman, J. E. Bowen, and O. I. Milner, *Anal. Chem.*, **31**, 2052 (1959).
(9) I. Hoffman and A. D. Gordan, *J. Assoc. Offic. Agric. Chem.*, **47**, 629 (1964).

RECEIVED for review February 2, 1976. Accepted June 7, 1976. This work is taken in part from the Masters Thesis submitted by Gerald Kellen to the Graduate School of Loyola University of Chicago.

Generalized Spectral Decomposition Method Applied to Infrared, Ultraviolet, and Atomic Emission Spectrometry

Harvey S. Gold, Carl E. Rechsteiner, and Richard P. Buck*

William R. Kenan, Jr., Laboratories of Chemistry, University of North Carolina, Chapel Hill, N.C. 27514

A generalized algorithm for symmetric and asymmetric Gaussian decomposition of spectral data is presented. Quantitative and qualitative applications in infrared, ultraviolet, and atomic emission spectroscopy are outlined in terms of general utility. The limitations and scope of curve fitting in general, and with regard to the current algorithm, are considered. For frequently encountered situations, use of Gaussian and bi-Gaussian parameters is shown to result in sufficient computational accuracy. For the first time, computerized decomposition is used to improve the apparent resolution in atomic spectroscopy. Decomposition of weak interligand transfer bands in near infrared spectra of ruthenium dimers is reported. Electronic spectra of a series of ruthenium bis(bipy) complexes have been decomposed and interpreted.

Present curve fitting methods are generally written specifically for one particular type of spectrum and yield information regarding the area or height of components of a larger envelope. Numerous studies have been made of the optimum mathematical model or procedure to be used in a given situation. Publications describing digital computer decomposition techniques are especially abundant in the area of infrared spectroscopy (*1–6*). Fields such as uv-visible absorption spectroscopy have also received wide attention, but other areas, notably those that involve emission rather than absorption processes, have been largely ignored. While the various specialized techniques function well in theory, the high specificity inherent in their algorithms required the user to have rather detailed knowledge of the expected outcome of the decomposition process. At best, several different algorithms, techniques, and/or programs are required to encompass the range of spectra commonly encountered in a laboratory situation.

In a significant study, Vandeginste and De Galan (*7*) have completed a quantitative study of the effect of several sources of error that can arise in curve fitting procedures. This rigorous analysis follows the more quantitative investigations of Beacham and Andrew (*8*) on the number of simultaneous solutions to be expected from a decomposition algorithm. Perram (*9*) and Davis (*10*) concluded that numerical decomposition of a digitized contour is not necessarily unique. These conclusions are extremely significant and must be dealt with in a convincing fashion by all decomposition programs.

Perhaps the most important sources of error in curve fitting arise from inaccurate determination of the true number of bands present within a spectral envelope. This problem is exacerbated by input profiles which have only two inflection points. As the number of component bands increases, the possibility for ambiguous solutions increases. This is true even in those instances when the number of components is explicitly known. A further complication is due to the effect of various initial values upon the future course of the curve fitting algorithm. Schwartz (*11*) studied the consequences of poorly selected initial parameters upon the direction of solution, and found that either divergence or convergence (upon either correct or erroneous component characterization parameters) could result.

In the present study, an algorithm was developed based upon the concept that the resulting computer program should be capable of taking raw input data characteristic of a variety of spectral types and produce a decomposed spectrum that is meaningful for the spectral type under consideration. The result, SPECSOLV, is a FORTRAN program that requires a minimum of user input. When there has been a choice between ease of use and maximum computational efficiency, the former has generally been selected, consistent with the requirement that the program execute in a computer core region less than that automatically allocated for either an IBM FORTRAN G or FORTRAN H compiler. The most recent link edited version of SPECSOLV, level 2.8, operates in 96K (FORTRAN H) or 84K (FORTRAN G) of GO step core on either the TUCC (Triangle Universities Computation Center) IBM 370/165 or the UNCCC (University of North Carolina Computation Center) IBM 360/75.

Consistent with minimizing required user input, initial input values are derived by SPECSOLV from the experimental spectrum. Essentially the only input information required (other than a digitized spectrum) is a guess as to the number of components present. Other input parameters have suggested values and/or assigned default values. A possible shortcoming in curve fitting is ambiguity in locating the baseline. This problem is resolved in SPECSOLV by three different schemes. Analysis of many spectral types commonly involves the removal of a background or blank spectrum which is typically that of a solvent with a complex spectra in the area of interest. In this case, the blank is subtracted pointwise after appropriate interpolation of the input baseline spectrum over the explicit range. A second mode involves explicit definition of the baseline which is assumed constant over the entire range. This allows for shifted, uniform baselines and for analysis of small regions where baseline changes are minimal. The third mode depends upon a program of baseline calculations based upon the root height squared ordinate value.

SPECSOLV is designed to decompose a wide variety of spectral types. This paper details applications to ultraviolet/visible, atomic emission, and infrared spectroscopies. Future papers will deal with applications to chromatography, organic mass spectrometry, spark source spectrometry, nuclear magnetic resonance, atomic absorption, and molecular fluorescence spectroscopy.

DECOMPOSITION METHODS

Functional Forms. The large number of spectral types that can be analyzed by use of SPECSOLV has necessitated restricting the choice of line shapes for the component peaks. In most cases, selection of either a pure Gaussian or of an asymmetric Gaussian (hereafter referred to as a bi-Gaussian) is sufficient. While some spectra intrinsically possess a Cauchy (Lorentzian) shape rather than a Gaussian form, these seeming limitations can be ignored in many instances. Frequently, evaluation of reference and experimental spectra using the same characterization parameters results in effective cancellation of the particular band shape selected.

Studies by Fraser and Suzuki (*1*) and Jones (*12*) have shown for infrared bands in particular, that absorption bands are commonly combinations of Gaussian and Cauchy functions. Use of a product or summation gives rise to a large number of parameters that need to be implicitly evaluated. The sum function has the form

$$T(x) = \exp 2.3\{W_c A_c[1 + b_c^2(x - x_{max})^2]^{-1} + W_g A_g \exp[-b_g^2(x - x_{max})^2]\} = \exp 2.3\, A(x) \quad (1)$$

where $T(x)$ = transmittance, W_c = Cauchy weighting factor, A_c = peak height of the Cauchy function, x_{max} = peak position, b_c = twice the inverse full-width-half-maximum of the Cauchy function, W_g = Gaussian weighting factor, A_g = peak height of the Gaussian function, $b_g = 2(\ln 2)^{1/2}$ divided by the full-width-half-maximum of the Gaussian function, and $A(x)$ = absorbance.

Pitha (*13*) has demonstrated that parameter reduction in Equation 1 will result in little loss of decomposition accuracy. Vandeginste and De Galan (*7*) have used this strategy by assuming the fixed relationship that $b_g = 0.8\, b_c$. This assumption allows algorithms that employ pure Gaussian or pure Cauchy functions a measure of interconvertibility, although this is only approximate at best.

In this study, the components are assumed to be of Gaussian or bi-Gaussian form. A Gaussian distribution is of the form

$$Y = Y_0 \exp\{-\ln 2[2(X - X_0)/\Delta X_{1/2}]^2\} \quad (2)$$

where Y_0 is the peak height, X_0 is the peak position, and $\Delta X_{1/2}$ the full-width-half-maximum. Equation 2 can readily be seen to be identical to Equation 1 in the limit when $W_c = 0$ and $W_g = 1$. To allow for asymmetry, Gaussian fits can be obtained for the right and left sides of a single peak. The form of a bi-Gaussian then becomes

$$Y = A_g \exp\{-[\beta_g(-)]\,(X - X_{max})^2\} \quad \text{for } X \text{ less than } X_{max} \quad (3a)$$

$$Y = A_g \exp\{-[\beta_g(+)]\,(X - X_{max})^2\} \quad \text{for } X \text{ greater than } X_{max} \quad (3b)$$

when expressed in a form similar to Equation 1; $\beta_g(-)$ and $\beta_g(+)$ signify peak width parameters of the two halves of the peak. The distance between inflection points is $(2/\beta_g)^{1/2}$.

Mode of Calculation. Decomposition of curves is accomplished by use of a least squares routine. In general, a least squares procedure attempts to evaluate the parameters $P_j = P_1, P_2, P_3, \ldots$ which results in the best fit of a function $Y(x, P_j)$ to a discrete 2×2 matrix of y_i vs. x_i. When the sum of the squares of the deviations at each point on the curve is minimized, the optimum value of P_j has been determined. This may be described by

$$S = \sum_j [y(s_i, P_j) - y_i]^2 = \text{minimum} \quad (4)$$

In the traditional least squares procedure, $y(x, P_j)$ is typically a polynominal in x with coefficients P_j. The result is a set of simultaneous equations whose solution gives P_j. In the case of a function $y(x, P_j)$ which is a summation of Gaussian components, the system is a set of simultaneous equations nonlinear in P_j's. This precludes obtaining an exact solution and necessitates an iterative procedure.

Any iterative procedure requires initial parameters. In the case of a spectrum, the initial values are easily obtained. Ideally, the program algorithm would determine all initial values. Reality, however, makes this prohibitively complex and quite an expensive computational problem. We have chosen to input a guess as to how many component peaks are present, allowing SPECSOLV to calculate initial rough values for the peak positions, heights, and half-widths from the digitized spectrum. At present, self-imposed core restrictions set a limit of 10 on the number of peaks present, and a spectral region restricted to accommodate this. Ongoing development is progressing on an interactive Time Sharing Option (TSO) version that will allow internal changes in peak number until an increase in peak number has negligible effect upon the sum of squared residuals as a function of peak number and convergence rate. Frequently, physical reality and knowledge of the system place a logical restraint on the number of peaks that can be present.

As a realistic problem in programming, the iterative convergence scheme is not likely to be totally perfect. To allow for this, tolerance factors are evaluated. To a great extent, these reflect the degree of uncertainty associated in digitizing a spectrum and also serve to place an upper bound on the degree of noise present in the computed spectrum.

Table I. Input and Output Units for Each of the Various Spectral Types Supported by SPECSOLV

Type[a]	Input units	Output units
1, 2, 3, 9, 10, 11	Angstroms	Kilokaysers
4, 5	Arbitrary	Same units
6, 8, 14	Distance	Distance
7	Time	Time
12, 13	Angstroms	Angstroms

[a] Key: 1 = ultraviolet/visible. 2 = infrared. 3 = Raman. 4 = General Gaussian. 5 = General bi-Gaussian. 6 = Photographic (mass spectrometry, etc.). 7 = time based chromatography. 8 = distance based chromatography. 9 = electronic absorption. 10 = near infrared. 11 = far infrared. 12 = atomic emission/fluorescence. 13 = atomic absorption. 14 = nuclear magnetic resonance.

COMPONENT DETECTION AND EVALUATION

The input spectrum, a discrete function typically of wavelength, is first baseline adjusted and then transformed into a second discrete function. This is achieved by way of interpolation of points along a cubic segment passing through the nearest four input points. At this stage, the interpolated spectrum may be a function of the input units, or variously transformed to a different set of units as in the transformation from angstroms to kilokaysers for ultraviolet-visible spectra. The interpolation interval is variable, being primarily a function of the type of spectrum being analyzed, and secondarily a function of the range of the data points. This flexibility allows for analysis of spectra with varying ranges while assuring that meaningful point density is maintained.

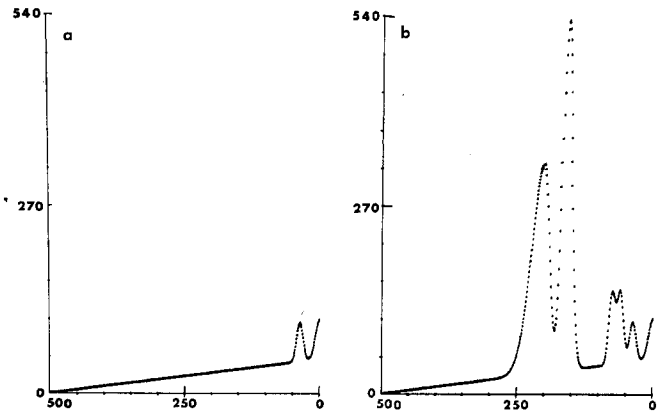

Figure 1. Synthetic test spectrum

(a) Baseline. (b) Baseline plus synthetic spectrum

A preliminary search for peaks is made, and initial peak characterization is achieved. The theoretical peaks calculated from these parameters are successively subtracted from the spectrum. An iterative procedure is begun which consists of adding, searching for, and subtracting each calculated peak from the residual spectrum. Following each iteration, the sum of squared residuals is compared with a tolerance factor. Iteration proceeds until the tolerance test is satisfied or until a user or program supplied iteration limit is reached.

At this point, final peak parameters consisting of peak height, width, and position are printed. In the case of electronic spectra, the oscillator strengths of the various transitions are also calculated.

Many options available to the user are not evident from the discussion above. Among these are selection of the type of spectral analysis (Table I), baseline modes, whether all iterations are to be displayed, input intervals, uncertainty in readings, path length (if appropriate), concentration (if appropriate), scaling factors, and output reference positions. Many of these have default values. In addition, plots can be obtained for input, calculated, and component peak spectra.

EXPERIMENTAL

Ultraviolet-visible spectra (*14*) were taken in Spectrograde acetonitrile using a Bausch and Lomb 210 spectrophotometer in the range from 850 to 220 nm. The ruthenium nitrosoarene complexes selected for analysis were synthesized by Bowden (*14*). The spectra were digitized at 100-Å intervals and subjected to bi-Gaussian analysis prior to interpretation.

Atomic emission spectra were run using a nitrous oxide–acetylene flame and an Instrumentation Laboratory, Inc. Model 334 auto scan module set to "slow". The output signal was interfaced to a Beckman 10-inch recorder via an Instrumentation Laboratory Model 331 power supply control module. The resultant spectra were digitized at 1.0-Å intervals in the range from 5877 to 5907 Å and subsequently decomposed.

Infrared spectra in the near ir range were recorded using a Cary 14 spectrophotometer. The ruthenium compounds used were synthesized by Powers et al. (*15*). The spectra were digitized at 100-Å intervals and analyzed by use of SPECSOLV.

Respective emission or absorbance measurements were made directly from the chart paper readout. The emission for atomic emission experiments was measured in arbitrary units, making use of appropriate scaling factors for attenuator adjustments.

RESULTS AND DISCUSSION

Synthetic Spectrum. A synthetic test spectrum was prepared by summing three symmetric and one highly asymmetric Gaussian peaks. The sum was superimposed on a sloping baseline and background peaks were added. The baseline and combination baseline/synthetic spectrum are shown as Figures 1a and 1b, respectively. This spectrum was decomposed by SPECSOLV. The decomposed components are depicted in Figure 2. The agreement between theoretical positions, half-widths, and heights, and those calculated by SPECSOLV is excellent.

Enhanced Atomic Emission Spectroscopy. A convenient definition of resolution in spectroscopy is

$$R = \lambda/\Delta\lambda$$

where λ is the wavelength and $\Delta\lambda$ is the minimum separation of two peaks of equal size which are adequately resolved. The easiest way of specifying quantitatively the adequacy of the separation is shown in Figure 3. For two peaks of equal size and width and of Gaussian line shape, the combined spectrum is

$$I = A\left\{\exp\left[\frac{-(\lambda-\lambda_1)^2}{2\sigma^2}\right] + \exp\left[\frac{-(\lambda-\lambda_2)^2}{2\sigma^2}\right]\right\} \quad (5)$$

At point λ, $I = \alpha A$ and $(\lambda - \lambda_1) = (\lambda - \lambda_2) = \Delta\lambda'$. Solving for $2\Delta\lambda'$ where $\Delta\lambda' = \Delta\lambda/2$ gives

$$2\Delta\lambda' = 2[2\ln(2/\alpha)]^{1/2}\sigma \quad (6)$$

The resolution is therefore

$$R = \lambda/\{2[2\ln(2/\alpha)]^{1/2}\sigma\}$$

which when $\alpha = 0.1$ gives

$$R = \lambda/4.895\sigma \quad (7)$$

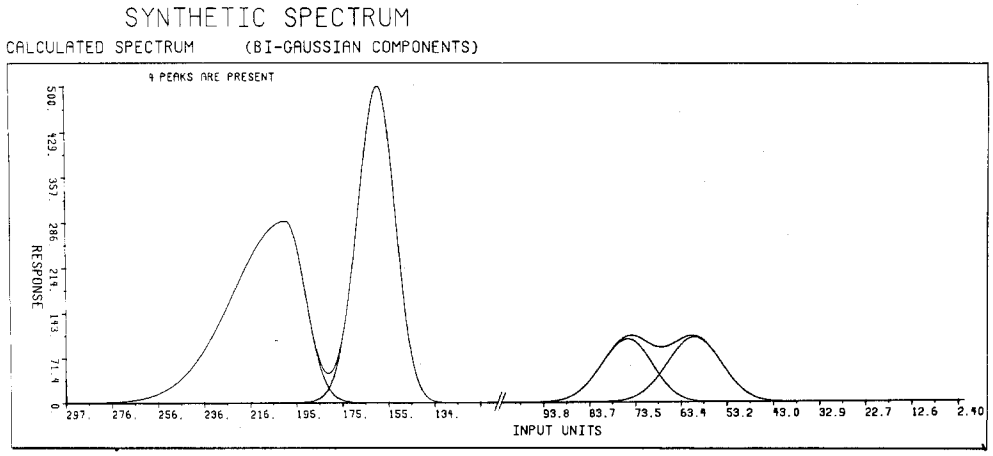

Figure 2. Decomposed synthetic test spectrum

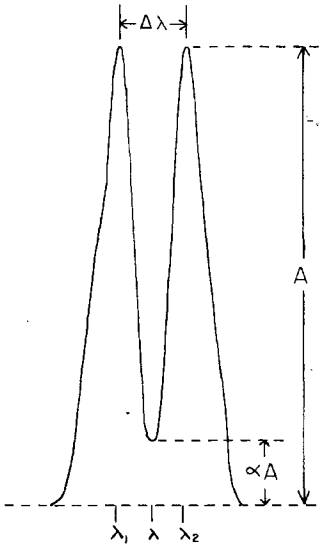

Figure 3. Parameters for calculating resolution using the 10% valley definition ($\alpha = 0.1$)

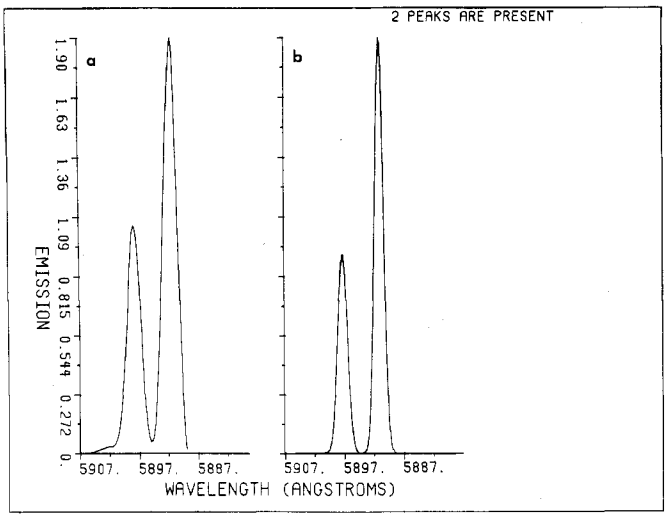

Figure 4. Atomic emission spectrum of the sodium D doublet with 20-micron entrance slits

(a) Original spectrum. (b) Decomposition spectrum with envelope and components shown

In this experiment, the sodium doublet at 5890 Å and 5896 Å was observed at different slit settings. The doublet made it easy to observe quantitatively the decrease in resolution as a function of increased slit width. This is apparent from Figures 4a, 5a, and 6a which show representative spectral envelopes that served as input data for decomposition. Increase in relative emission intensity is a measure of the increased throughput through the monochromator slits.

Quantitative analysis of the resolution has little meaning for slit widths of 80 microns or more, since α is in these cases clearly much larger than 0.1. However, use of SPECSOLV allows the apparent resolution to be tremendously enhanced, Figures 4b, 5b, and 6b depict the calculated spectral envelope as well as the component peaks. In many cases, the two are not distinguishable on the plotter output. To test for any asymmetry introduced by the instrument, the decomposition was performed in bi-Gaussian mode. As Table II shows, both bi-Gaussian half-width parameters are nearly equal, indicating highly symmetric peaks.

Bi-Gaussian half-widths can be converted into standard deviations by the relationship

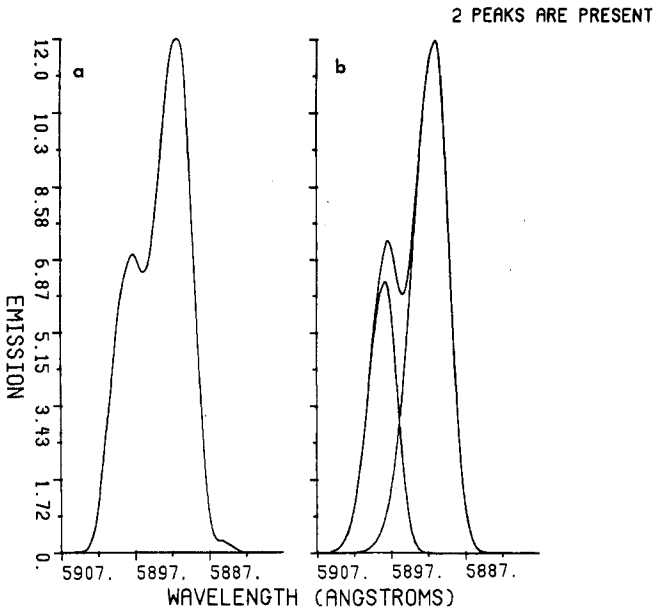

Figure 5. Atomic emission spectrum of the sodium D doublet with 160-micron slits. The result of deconvolution is quite apparent

(a) Original spectrum. (b) Decomposition spectrum with envelope and components shown.

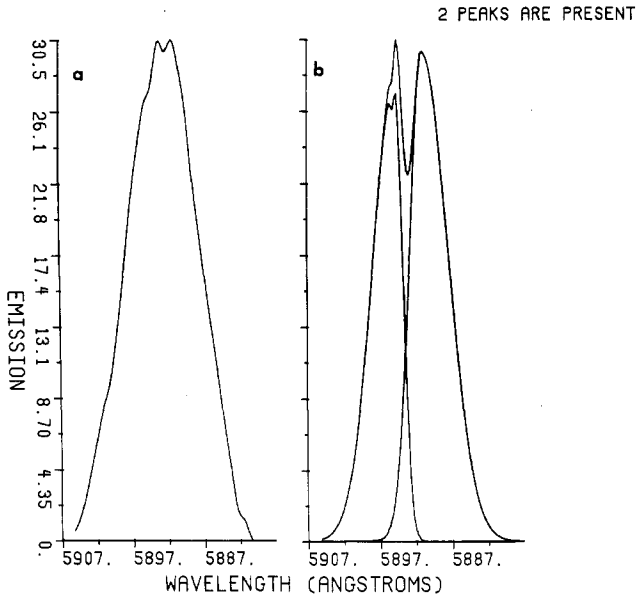

Figure 6. Atomic emission specrrum of the sodium D doublet with 320-micron slits. The original spectrum exhibits self-absorption due to the large slit widths. This is reflected in the failure of the program to converge. Nevertheless, the final peak values are useable

(a) Original spectrum. (b) Decomposition spectrum with envelope and components shown

$$\sigma = \Delta X_{1/2}/\{2[2\ln(2)]^{1/2}\} = \Delta X_{1/2}/2.254 \quad (8a)$$

where

$$\Delta X_{1/2} = [\Delta X_{1/2}(-) + \Delta X_{1/2}(+)]/2$$

In general, substitution of Equation 8a in Equation 7 shows that for any α

$$R = \lambda/(\Delta X_{1/2}\{1-[\ln(\alpha)/\ln(2)]\}) \quad (8b)$$

Use of Equation 7 and the standard deviations summarized

Table II. Peak Parameters Calculated by SPECSOLV for the Sodium D Doublet Using Various Entrance Slit Widths[a]

Slit width, microns	$\Delta X_{1/2}(+)$ (Å)		$\Delta X_{1/2}(-)$ (Å)		Std dev, σ (Å)		Calculated peak positions (Å)		Height peak 2/ height peak 1
	Peak 1	Peak 2	Peak 1	Peak 2	Peak 1	Peak 2	Peak 1	Peak 2	
20	0.636	0.889	0.636	0.889	0.270	0.378	5896.2	5889.9	0.6165
40	0.945	0.852	0.945	0.852	0.401	0.362	5897.1	5890.8	0.4798
80	1.31	2.19	1.31	2.19	0.556	0.930	5896.1	5889.9	0.5405
160	2.25	1.89	3.44	2.67	1.21	0.968	5897.6	5890.8	0.5216
320	4.19	3.74	2.05	1.28	1.33	1.07	5895.2	5890.3	0.9273

[a] The peak referred to as 1 is due to the $^2P_{1/2}$ transition, while the 2 peak is due to the $^2P_{3/2}$ transition. The data for the 3.20-micron slit width failed to converge upon analysis; the results reported are the values calculated at the time that the program iteration limit was reached. $\Delta X_{1/2}(+)$ and $\Delta X_{1/2}(-)$ are peak half-widths for each side of a peak.

Table III. Resolution and Grating Factors for Atomic Emission Spectroscopy for Various Slit Widths after Decomposition by SPECSOLV

Slit width, microns	Resolution factor (at 589 Å)	Active grating width, mm
20	4460.4	0.26
40	3003.4	0.39
80	2166.1	0.54
160	995.3	1.19
320	905.5	1.30

Table IV. Results of the Decomposition of the Visible Spectra of $Ru(bipy)_2(N(O)C_6H_4R)X^+$ Showing Component Positions and Extinction Coefficients

	λ, nm	$\bar{\nu}_{max}$(kK)	ϵ
X = NO_2; R = $N(CH_3)_2$	580	17.24	3.2×10^4
	420	23.76	2.2×10^4
	478	20.86	5.9×10^3
	610	16.39	4.7×10^3
X = NO_2; R = $NHCH_3$	547	17.44	2.1×10^4
	414	24.22	1.9×10^4
	480	20.79	5.3×10^3
X = Cl; R = $N(CH_3)_2$	600	16.67	2.7×10^4
	424	23.56	2.0×10^4
	482	20.75	5.4×10^3
	687	14.54	2.4×10^3
X = Cl; R = $NHCH_3$	577	17.33	1.9×10^4
	411	24.41	1.4×10^4
	478	20.93	7.2×10^3
	672	14.79	2.6×10^3
X = Cl; R = OH	572	17.50	1.6×10^4
	464	21.55	7.6×10^3
X = Cl; R = CH_3	453	22.06	7.6×10^3
	521	19.19	1.9×10^3
	840	11.90	3.4×10^2
	418	23.93	6.4×10^2
X = Cl; R = H	425	23.45	7.8×10^3
	505	19.80	1.0×10^3
	691	14.46	2.1×10^2
X = Cl; R = Br	432	23.16	6.7×10^3
	506	19.77	1.7×10^3
	537	18.57	4.8×10^2
X = Cl; R = NO_2	418	23.95	$\sim 7.8 \times 10^3$
	511	18.52	$\sim 1.0 \times 10^3$

in Table II gives a resolution index which appears in Table III for each slit width.

The resolution clearly increases as the slit width is decreased, but the magnitude of the effect diminishes as the slits become narrower. This is attributed to optical aberrations in the collimating mirror and to Fraunhoffer diffraction at the slit edges which makes the slit width appear to be larger.

The theoretical equation for resolution can be stated as

$$R = n \times N \tag{9}$$

where n = spectral order and N = gooves per mm divided by the active grating width. R as calculated from Equation 9 is a hypothetical resolution for zero slit width, yet finite grating illumination. Since the spectral order is 1 at 5890 Å for a grating blazed at 3500 Å, the active grating width for each slit width can be easily evaluated, with results as shown in Table III. The increase in active grating width explains the higher relative emission throughput as slit width increases.

A significant result of this analysis is the demonstration that a very poorly resolved atomic emission spectrum can be quantitatively resolved into component peaks. While the accuracy of the calculated positions for the $^2P_{3/2}$ and $^2P_{1/2}$ transitions is not perfect, the deviation from accepted positions (5890 and 5896 Å, respectively) is generally well within experimental accuracy. It is significant that the ratio height of the peaks is relatively constant (Table III), and in accord with the generally accepted ratio of quantum efficiencies for the two transitions.

Electronic Spectral Decomposition. To test the capabilities of the decomposition abilities of SPECSOLV, the decision was made to analyze the hitherto unexplained electronic spectra of various ruthenium(II) nitrosoarene complexes. These visible and ultraviolet spectra exhibit broad bands and asymmetric peaks that suggest overlapping absorption bands. The resolved spectra and component peaks of $Ru(bipy)_2(N(O)C_6H_4N(CH_3)_2Cl^-PF_6^+$ and $Ru(bipy)_2(N(O)C_6H_4CH_3)Cl^-PF_6^+$ are reproduced as Figures 7 and 8 respectively.

In addition to these compounds, other substituted ruthenium nitrosoarene compounds were analyzed. The results of spectral decomposition are summarized in Table IV. Using these results, Bowden (*14*) has shown that the positions of the absorption bands follow a predictable pattern and that band transition assignments can be made with relative certainty. Reference to Figures 7 and 8 indicate that many of the smaller component bands cannot be identified, let alone characterized, before decomposition. However, following decomposition, even relatively small shifts in band position and strength are readily apparent and correlate quite well with changes at the metal center and at remote ligand sites.

In addition to band parameters, SPECSOLV analysis of electronic spectra is designed to determine the oscillator strength of each transition. A discussion of the nature and basis of absorption bands will serve to rationalize calculation of the oscillator strength and to justify assumption of a Gaussian line shape. In 1930, Kuhn and Braun (*16*) proposed a formulation for the Gaussian function in wavenumber σ as

$$\epsilon = \epsilon_0 \exp[-(\sigma - \sigma_0)^2/\nu^2] = \epsilon_0\, 2^{-(\sigma-\sigma_0)^2/\delta^2} \tag{10}$$

with ϵ_0 of the maximum at σ_0 and the half width δ(since $\epsilon = \epsilon_0/2$ for $\sigma = \sigma_0 \pm \delta$) being equal to $\nu\sqrt{\ln 2}$. The area of a band in this representation is proportional to the product of ϵ_0 and δ, since for $\sigma \gg \delta$

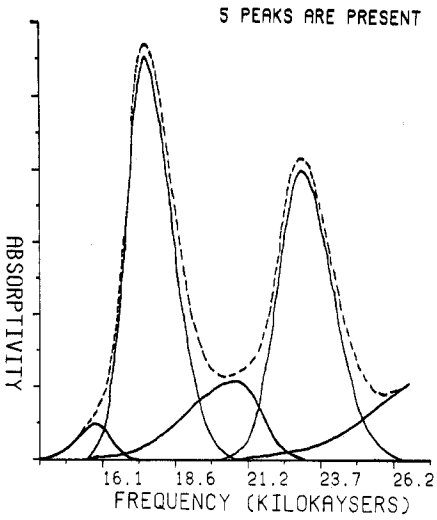

Figure 7. The ultraviolet spectrum of $Ru(bipy)_2(N(O)C_6H_4N(CH_3)_2)$-$Cl^-PF_6^+$ before and after bi-Gaussian decomposition by SPECSOLV. The dotted line (- - -) indicates the original spectral envelope after digitization

$$\int_0^\infty \epsilon d\sigma = 2\epsilon_0 \int_0^\infty \exp[-(\sigma - \sigma_0)^2 \ln(2/\delta^2)] d(\sigma - \sigma_0)$$

$$= \epsilon_0 \delta (\pi)^{1/2} (\ln 2)^{-1/2} = 2.11289\ \epsilon_0 \delta \quad (11)$$

This area is proportional to the oscillator strength, P, of the transition, since

$$P = \frac{1000\ mc^2}{N\ \pi e^2} 2.30 \int \epsilon d\sigma = 4.32 \times 10^{-9} \int \epsilon d\sigma$$

$$= 9.20 \times 10^{-9}\ \epsilon_0 \delta \quad (12)$$

where N is Avogadro's number, e is the charge of the electron, m is the mass of the electron, and c is the velocity of light.

The Franck–Condon principle functions as a projection of the vibrational wave function of the electronic ground state on the potential curve of the excited electronic state. Since the square of the vibrational wave function of the lowest level of a harmonic oscillator is a Gaussian curve with a half width $x_c(\ln 2)^{1/2}$, the use of a pure Gaussian or bi-Gaussian form to describe electronic spectral band shapes is substantiated.

This allows spectral analysis to be focused on only three parameters: The position, the half-width, and the area or intensity.

1) The position indicates the energy difference between the excited and the ground state at the minimum of the potential curve of the ground state.

2) The half-width is a measure of the anti-bonding character of the excited state. The anti-bonding character is a function of the internuclear distance, r. While an estimate of the increase in r from b_g alone without detailed knowledge of the force constant of the excited ground state (and its anharmonicity), is not possible in depth, a corresponding increase in r with increasing b_g within a given complex is found.

3) The area of intensity is proportional to $e_0 b_g$. An extremely valuable consequence of the Gaussian shape is that independent bands can be identified from the presence of shoulders or even from the presence of weak inflections.

The utility of the theoretical results is apparent from inspection of aspects of Figures 7 and 8. In the former, note the detection of the small peak at 16.39 kilokaysers. The latter figure shows two peaks at 19.19 and 23.93 kilokaysers, whose presence, but not shape, is only hinted at in the original spectrum. Note the small, highly asymmetric peak at 11.90 kilokaysers. The positions and oscillator strengths of these component peaks are properly correlated with theoretical predictions (*14*).

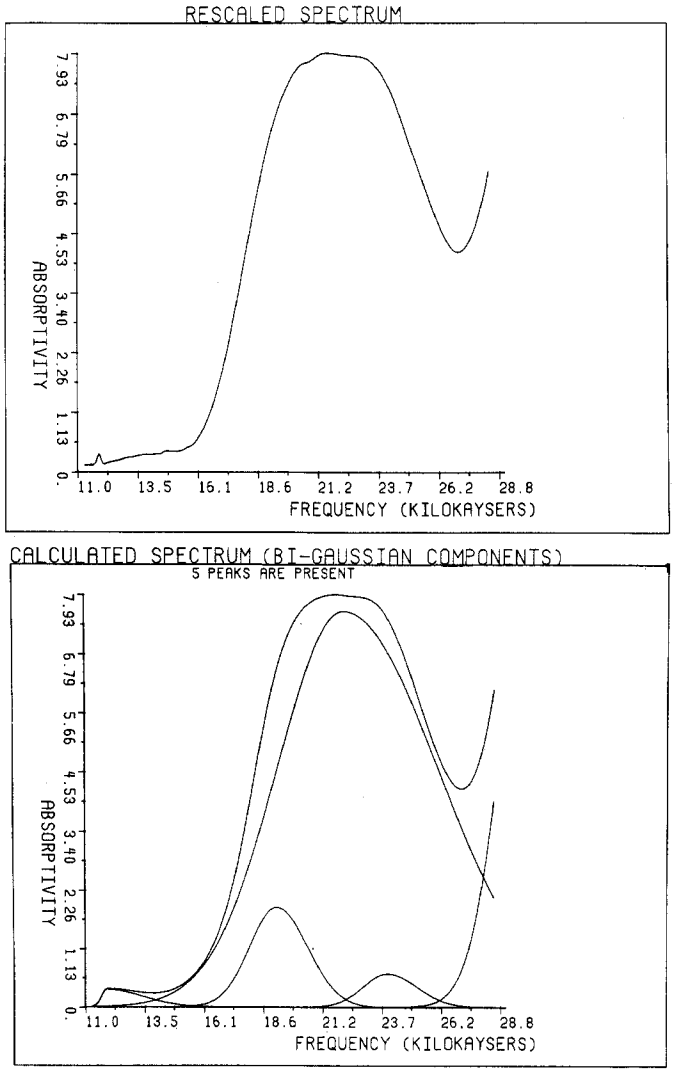

Figure 8. The ultraviolet spectrum of $Ru(bipy)_2(N(O)C_6H_4CH_3)Cl^-PF_6^+$

(a) The original spectral envelope after digitization. (b) Component peaks and calculated envelope after bi-Gaussian decomposition

Infrared Spectral Analysis. In this area of investigation, a study of ruthenium dimer complexes was undertaken (*15*). This area of interest involves metal centers bridged by ligands with the metals having weak interactions across the bridge. In the case of mixed valence compounds, intervalence transfer bands appear in the near infrared. These transitions are quite weak, typically having molar extinction coefficients of 100 to 500, and usually appear as shoulders or lie very close to intense bands near 5000 Å. The problem is therefore to detect solvent dependent shifts of a small band located in the vicinity of a much stronger band. The results of this study have recently been presented by Powers et al. (*15*).

Figure 9 presents the successful decomposition of the near infrared spectrum of the PF_6 salt of $[(NH_3)_5Ru(pyz)Ru$-$Cl(bipy)_2]^{4+}$ (pyz = pyrazine). It shows a small band adjacent to a much larger electronic absorption band, only part of which is shown. Characterization of similar compounds in a variety of solvents have been successfully accomplished by use of SPECSOLV, whereas any such analysis was heretofore impossible.

CONCLUSIONS

In this paper, we have presented three specific applications of the decomposition algorithm of SPECSOLV. We cannot overemphasize the simplicity of using this particular program. While there has admittedly been some loss of accuracy by

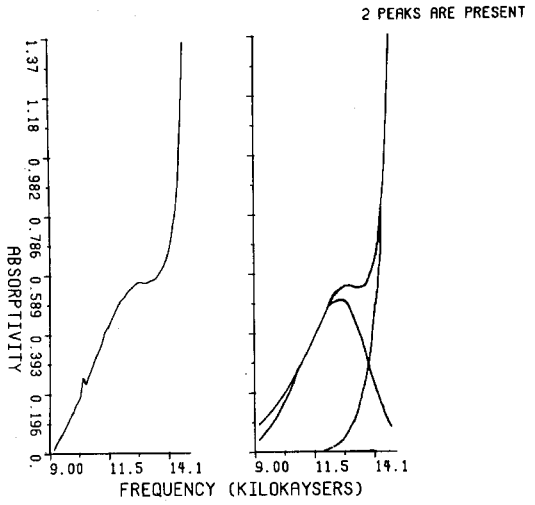

Figure 9. The near infrared spectrum of the PF_6 salt of $[(NH_3)_5Ru(pyz)\text{-}RuCl(bipy)_2]^{3+}$ showing a weak intervalence transfer band as a shoulder to a much stronger absorption. The larger component is only partially shown

assuming purely Gaussian or bi-Gaussian line shape, the resulting ease and simplicity of use is significant. SPECSOLV has been used successfully by numerous researchers in this department who lack any knowledge of decomposition schemes or computerized methods. We feel that the ability to require as input only information immediately available from the experiment is a significant advance in this field. Many of the possible input parameters to SPECSOLV have default values. In any event, only the choice of the number of peaks to be found affects the search outcome; other parameters largely deal only with output and input forms. While the selection of the number of component peaks may seem arbitrary, underestimation is frequently signaled by a failure to converge after iteration, while overestimation gives rise to small "noise" peaks.

SPECSOLV is available for public user distribution through the Triangle Universities Computation Center network. Inquires regarding availability of source listings or decks may be addressed to: Program Librarian, University of North Carolina Computation Center, Phillips Hall, Chapel Hill, N.C. 27514.

ACKNOWLEDGMENT

The authors acknowledge the work of W. L. Bowden and M. J. Powers, whose research problems have provided applications tests for SPECSOLV. The various user oriented suggestions of J. A. Baumann have been quite valuable. One of us (HSG) thanks W. E. Hatfield for his initial inspiration.

LITERATURE CITED

(1) R. D. B. Fraser and E. Suzuki, *Anal. Chem.*, **41**, 37 (1969).
(2) K. S. Seshadri and R. N. Jones, *Spectrochim. Acta., Part A*, **19**, 1013 (1963).
(3) R. P. Young and R. N. Jones, *Chem. Rev.*, **71**, 219 (1971).
(4) A. M. Kabiel and C. H. Boutros, *Appl. Spectrosc.*, **22**, 121 (1968).
(5) F. C. Strong, *Appl. Spectrosc.*, **23**, 593 (1963).
(6) J. Pitha and R. N. Jones, *Can. J. Chem.*, **44**, 3031 (1966).
(7) B. G. M. Vandeginste and L. De Galan, *Anal. Chem.*, **47**, 2124 (1975).
(8) J. R. Beacham and K. L. Andrew, *J. Opt. Soc. Am.*, **61**, 231 (1971).
(9) J. W. Perram, *J. Chem. Phys.*, **49**, 4245 (1968).
(10) A. R. Davis et al., *Appl. Spectrosc.*, **26**, 384 (1972).
(11) L. M. Schwartz, *Anal. Chem.*, **43**, 1336 (1971).
(12) R. N. Jones, *Pure Appl. Chem.*, **18**, 303 (1969).
(13) J. Pitha and R. N. Jones, *Can. J. Chem.*, **45**, 2347 (1967).
(14) W. L. Bowden, Doctoral Thesis, University of North Carolina, Chapel Hill, N.C., 1975.
(15) M. J. Powers, R. W. Callahan, D. J. Salmon, and T. J. Meyer, in press, 1976.
(16) W. Kuhn and E. Braun, *Z. Phys. Chem., Abt. B*, **8**, 281 (1930); **9**, 426 (1930).

RECEIVED for review March 1, 1976. Accepted May 13, 1976. This work was supported by a grant from the National Science Foundation, MPS75-00970.

Determination of Relative p*K* Values of Dibasic Protolytes by Regression Analysis of Absorbance Diagrams

Friedrich Göbber* and Jürgen Polster

Institute for Physical and Theoretical Chemistry, University of Tübingen, D-7400 Tübingen 1, Federal Republic of Germany

The absorbances of a dibasic acid measured at two distinct wavelengths with varying pH as the parameter form a conic section which can be determined by an appropriate least-squares method. The conic section and the absorbances of the acid in its nondissociated form and its completely dissociated form allow for the determination of the difference of the p*K* values (ΔpK) and the absorbances of the ampholyte without using pH values. If data from more than two wavelengths are available, one ΔpK value can be computed for each pair of wavelengths. Because the reliabilities of the particular results are considerably different, an extensive error analysis is necessary to obtain an averaged ΔpK value together with an estimated standard deviation. The results for *o*-phthalic acid are given as an example.

The evaluation of the measurements from multistep acid–base equilibria requires the pH values in order to obtain the p*K* values and the absorbances of intermediates (*1–3*). Because the pH measurements are relatively inaccurate compared to spectrophotometric data, we investigated what information can be extracted from the more reliable spectrophotometric data alone. We found that, in the case of dibasic acids, the difference between the p*K* values and the absorbance of the ampholyte can be obtained without using pH measurements (*4, 5*).

Our former method involves arbitrariness to some extent. It needs a tangent to a curve which is only represented by some data points. This can be done by hand or by fitting a low order polynomial to a subset of the data points. In either case, certain decisions cannot be derived from exact criteria. In the present paper, we describe how the results can be obtained from the data by unique numerical procedures. These are executed by means of a computer program which allows for the evaluation of the statistics of the data, thereby providing estimates for the reliability of the results. This FORTRAN program is available from the authors.