

spectral similarity, either prior or after interpretation; comparing spectral generation techniques to identify an optimal technique for unknown organic mixtures; and identifying the limits of detection for this technique by preparing mixtures at varying concentrations.

These improvements will be followed by comparing this method to presently available analysis techniques by analyzing actual organic hazardous waste samples.

### ACKNOWLEDGMENT

The authors wish to express their sincere appreciation to Sterling Tomellini for his help in adapting INTERP to our Nicolet 1280 computer system.

**Registry No.** Acetone, 67-64-1; anthracene, 120-12-7; benzene, 71-43-2; 2-butanone, 78-93-3; chlorobenzene, 108-90-7; chloroform, 67-66-3; *o*-cresol, 95-48-7; *p*-cresol, 106-44-5; dibutyl phthalate, 84-74-2; *o*-dichlorobenzene, 95-50-1; *m*-dichlorobenzene, 541-73-1; 1,1-dichloroethane, 75-34-3; 1,1-dichloroethene, 75-35-4; dichloromethane, 75-09-2; 1,2-dichloropropane, 78-87-5; 2,4-dimethylphenol, 105-67-9; ethylbenzene, 100-41-4; hexachlorocyclopentadiene, 77-47-4; 2-hexanone, 591-78-6; 4-methyl-2-pentanone, 108-10-1; pentachloroethane, 76-01-7; phenol, 108-95-2; styrene, 100-42-5; 1,1,2,2-tetrachloroethane, 79-34-5; toluene, 108-88-3; 1,2,4-trichlorobenzene, 120-82-1; 1,1,1-trichloroethane, 71-55-6; 1,1,2-trichloroethane, 79-00-5; 1,2,3-trichloropropane, 96-18-4; *o*-xylene, 95-47-6.

### LITERATURE CITED

- (1) "Resource Conservation and Recovery Act of 1976"; Public Law 94-580, 1976.
- (2) The Hazardous and Solid Waste Amendments of 1984 *Congr. Rec.* 1984, Oct 3, H11103.
- (3) "Comprehensive Environmental Response Compensation and Liability Act of 1980"; Public Law 96-510, 1980.
- (4) Puskar, M. A.; Levine, S. P.; Turpin, R. In "Protecting Personnel at Hazardous Waste Sites"; Levine and Martin, Eds.; Butterworths/Ann Arbor: Woburn, MA, 1985; Chapter 6.
- (5) Gurka, D. F. "Project Summary: Interlaboratory Comparison Study: Methods for Volatile and Semivolatile Compounds"; Environmental Monitoring Systems Laboratory, Las Vegas, NV, EPA-600/S4-84-027, June 1984.
- (6) Hallstedt, P. A.; Levine, S. P.; Puskar, M. A., submitted for publication in *J. Hazard. Waste Hazard. Mater.*
- (7) Frankel, D. S. *Anal. Chem.* 1984, 56, 1011.
- (8) Levine, S. P.; Puskar, M. A. *Am. Ind. Hyg. Assoc. J.* 1985, 46 (4), 181-186.
- (9) Gurka, D. F.; Betowski, L. D. *Anal. Chem.* 1982, 54, 1819-1824.
- (10) Wilkins, C. L.; Gliss, G. N.; White, R. L.; Brissey, G. M.; Onyiruka, E. C. *Anal. Chem.* 1982, 54, 2260-2264.
- (11) Shafer, K. H.; Hayes, T. L.; Brasch, J. W.; Jakobsen, R. J. *Anal. Chem.* 1984, 56, 237-240.
- (12) Puskar, M. A.; Levine, S. P. "Characterization of Bulk Materials on Remedial Action Sites: A Preliminary Comparison of Compatibility Testing, FT-IR/ATR and GC/MS", Proceedings of the National Conference on Hazardous Waste and Environmental Emergencies, May 1985.
- (13) Shafer, K. H.; Cooke, M.; DeRoos, F.; Jakobsen, R. J.; Rosario, O.; Mullik, J. D. *Appl. Spectrosc.* 1981, 35, 469.
- (14) Erickson, M. D. *Appl. Spectrosc.* 1981, 35, 181-184.
- (15) Griffiths, P. R.; Azarraga, L. V.; de Haseth, J.; Hannah, R. W.; Jakobsen, R. J.; Ennis, M. M. *Appl. Spectrosc.* 1979, 33, 543.
- (16) Lowry, S. R.; Huppler, D. A. Paper presented at 31st Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy, Atlantic City, NJ, March 10-14, 1980.
- (17) Rasmussen, G. T.; Isenhour, T. L.; Lowry, S. R.; Ritter, G. L. *Anal. Chim. Acta* 1978, 103, 213-221.
- (18) Antoon, M. K.; D'Esposito, L.; Koenig, J. L. *Appl. Spectrosc.* 1979, 33, 351-357.
- (19) Brown, C. W.; Lynch, P. F.; Obremski, R. J.; Lavery, D. S. *Anal. Chem.* 1982, 54, 1472-1479.
- (20) Maris, M. A.; Brown, C. W.; Lavery, D. S. *Anal. Chem.* 1983, 55, 1694-1703.
- (21) Osten, D. W.; Kowalski, B. R. *Anal. Chem.* 1985, 57, 908-917.
- (22) Frankel, D. S. *Anal. Chem.* 1984, 56, 1011-1014.
- (23) Trulson, M. O.; Munk, M. E. *Anal. Chem.* 1983, 55, 2137-2142.
- (24) Blaffert, T. *Anal. Chim. Acta* 1984, 161, 135-148.
- (25) Blackman, W. C.; Garnas, R. L.; Preston, J. E.; Swibas, C. M. "Chemical Composition of Drum Samples from Hazardous Waste Sites"; Management of Uncontrolled Hazardous Waste Sites, Nov 1984.
- (26) Eckel, W. P.; Trees, D. P.; Kovell, S. P. "Distribution and Concentration of Chemicals and Toxic Materials Found at Hazardous Waste Dump Sites"; Proceedings of the National Conference on Hazardous Waste and Environmental Emergencies, May 1985.
- (27) Woodruff, H. B.; Munk, M. E. *J. Org. Chem.* 1977, 42, 1761-1767.
- (28) Woodruff, H. B.; Munk, M. E. *Anal. Chim. Acta* 1977, 95, 13-23.
- (29) Woodruff, H. B.; Smith, G. E. *Anal. Chem.* 1980, 52, 2321-2327.
- (30) Woodruff, H. B.; Smith, G. E. *Anal. Chim. Acta* 1981, 133, 545-553.
- (31) Tomellini, S. A.; Saperstein, D. D.; Stevenson, J. M.; Smith, G. M.; Woodruff, H. B. *Anal. Chem.* 1981, 53, 2367-2369.
- (32) Tomellini, S. A.; Stevenson, J. M.; Woodruff, H. B. *Anal. Chem.* 1984, 56, 67-70.
- (33) Tomellini, S. A.; Hartwick, R. A.; Stevenson, J. M.; Woodruff, H. B. *Anal. Chim. Acta* 1984, 162, 227-240.

RECEIVED for review August 20, 1985. Accepted December 23, 1985. This work was supported by Grant 1-R01-OH02066-01 from the National Institute for Occupational Safety and Health of Centers for Disease Control.

## Fast Algorithm for the Resolution of Spectra

Richard A. Caruana, Roger B. Searle, Thomas Heller, and Saul I. Shupack\*

Departments of Chemistry and Computer Science, Villanova University, Villanova, Pennsylvania 19085

An algorithm is presented that resolves spectra into Gaussian bands by iteratively applying an efficient linear least-squares analysis to the individual bands. The advantages of the method are that the initial parameterization of the spectrum is often simpler and the procedure is very fast thereby enabling practical implementation on microcomputers. Tests performed on microcomputers indicate that the algorithm is capable of resolving difficult spectra at least 5-10 times faster than the conventional error space minimization techniques and can resolve simpler spectra about 50-100 times faster. Some of the techniques employed may be of benefit to users of other resolution methods. Novel measures of the goodness of fit are presented, which aid interpretation of the convergence path and the final fit obtained.

Methods have been developed to resolve spectra (or other data sets of overlapped bands) into their component bands

to obtain physically meaningful interpretations or to compact the information. Characteristically these methods present trade-offs in accuracy, speed, computational complexity, and power.

Computers with a significant amount of memory and processing power allow the use of nonlinear least-squares techniques, iterative methods that attempt to improve upon an initial set of parameter estimates by direct minimization of some measure of error. Most commonly employed are the gradient methods (1-4), simplex methods (5), and the many variations of these (1, 2, 6, 7). While many of these can be adapted to microcomputers, the amount of memory and/or time required for the analysis increases significantly with increasing problem size, making their use on microcomputers impractical for all but the simplest problems. Some of these methods also suffer from precision problems, and the use of double precision further aggravates the memory and time problems.

Several methods in the literature resolve spectra by performing transformations on the spectra that emphasize the separation of bands and aid in estimating band parameters (10, 11). Typically, these estimate band parameters by fitting the transformed spectrum via regression in regions where bands are relatively nonoverlapped and by utilizing derivatives. These approaches suffer from two major problems. First, they are usually noniterative and thus are very sensitive to the regions used to estimate parameters. If the spectrum is too overlapped no adequately unperturbed regions of the transformed spectrum may occur. Second, derivative methods are sensitive to noise. Digital filters can be used to minimize noise but can distort band shapes.

We present herein an algorithm that avoids these difficulties.

### METHOD

An isolated Gaussian band is characterized by its mean, area, and standard deviation. Plotting  $\log_e y$  vs.  $x$  for Gaussian data yields a quadratic in  $x$ :

$$\log_e y = a + bx + cx^2 \quad (1)$$

where

$$a = \log_e [\text{area}/(\text{std dev}(2\pi)^{1/2})] - \frac{1}{2} \text{mean}^2/\text{std dev}^2 \quad (2)$$

$$b = \text{mean}/\text{std dev}^2 \quad (3)$$

$$c = 1/(-2 \times \text{std dev}^2) \quad (4)$$

A least-squares fit to the quadratic yields estimates for  $a$ ,  $b$ , and  $c$  from which the three parameters of the original Gaussian band may be obtained:

$$\text{std dev} = 1/(-2c)^{1/2} \quad (5)$$

$$\text{mean} = b/(-2c) \quad (6)$$

$$\text{area} = (-1\pi/c)^{1/2} \exp(a - b^2/c/4) \quad (7)$$

This transformation and least-squares fit provides a fast (i.e., noniterative) means of determining the parameters of an isolated Gaussian band in the absence of noise. In the presence of noise or failure of the Gaussian model, estimates of the parameters of the best-fit Gaussian are obtained. These estimates are not the true best-fit parameters in the least-squares sense, but, in the absence of very large noise or large deviations from Gaussianity, this method provides reliable estimates and requires no further optimization. The use of the log transformation performs a weighting that minimizes the effects of noise where the spectrum is highest. This is desirable because the noise in a spectrum is usually greatest at higher absorbances (2). Another advantage of this method is that data need not be evenly spaced.

For a spectrum with relatively little overlap this technique may be employed by transforming and fitting the regions where each band clearly predominates. As the overlap in the spectrum increases the accuracy of the fit obtained by this noniterative procedure rapidly decreases.

The parameters for a band that is significantly influenced by other unknown bands cannot be directly determined. If, however, reasonable estimates of the parameters of the perturbing bands are known, some correction can be made by subtracting their contribution from the spectrum envelope, thus leaving the remaining band less perturbed. A fit performed on this partially corrected band should yield better estimates than those obtained from fitting the uncorrected band.

The heart of our algorithm is this process of correcting regions of the spectrum for the contribution from other bands via estimates of their parameters and then fitting these corrected regions by the noniterative procedure described pre-

viously. If reasonable estimates for all of the bands in the spectrum can be obtained, then better estimates can be obtained by performing this process for each band. This process is repeated, each iteration using the estimates obtained from the previous iteration to correct the spectrum, until convergence of the estimates of the parameters of all of the bands in the spectrum is obtained.

Partitioning the spectrum into regions where each band predominates is accomplished by pairs of data limits we call walls. Walls define the regions within which the spectrum will be corrected and fit. If there are  $K$  bands in the spectrum there are  $K$  wall pairs encompassing regions where each band is hopefully least perturbed. In these regions the estimates obtained from the fit will be the most reliable, especially in the earlier iterations where the corrections for the other bands are likely to be relatively poor. Walls need not necessarily bracket the band maximum, though they often will, since the band maximum is usually in the region where a band's effect on the envelope is most significant.

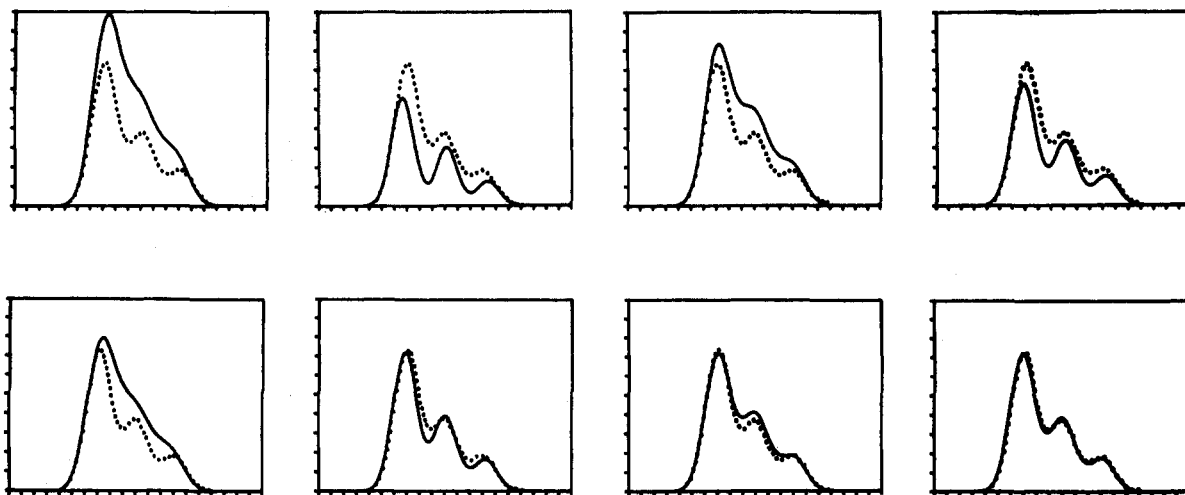
This simple algorithm may exhibit a "ringing" in the area estimates. Neighboring bands can only add apparent area to a band they perturb; therefore, it is expected that the first areas of all bands will be overestimated. If the estimated areas are too large and if the estimates for the standard deviations and means are roughly correct, then the contributions subtracted during the next iteration will be too large, causing the area estimates from the following iteration to be too small. This in turn will cause the area estimates of the next iteration to be too large and so on. Although the errors in the area estimates of successive iterations usually get smaller and convergence to the correct areas is eventually obtained, it is desirable to reduce this ringing.

Area renormalization is a technique we employ to damp the ringing in the area estimates. Area renormalization forces the total area predicted by the parameter estimates in the regions defined by the walls to equal the corresponding area measured directly from the spectrum. This is done by summing the area contributions predicted by the current parameter estimates in all of the walled regions. These predicted areas are calculated by an approximation formula for the definite integral of the normal distribution (8). Each band's walled area estimate is then corrected by dividing by the ratio of the sum of all predicted walled areas to the total walled area in the original spectrum. This second sum of areas is found by numerical integration using the trapezoid rule. The predicted area estimates are therefore in the same proportions as prior to renormalization.

### IMPLEMENTATION

Interactive menu-driven programs have been written in Pascal and run on Apple II, CP/M Z-80 based, and IBM PC microcomputers. The user programs provide considerable flexibility in initial parameterization. The user may specify from as little as the locations of the walls to a full initial parameterization of the spectrum. Initial parameter estimates for bands not explicitly parameterized are found automatically by applying the transformation to the walled regions after correcting the spectrum for any parameterized bands. This is quite convenient, and we have found that for many spectra it is sufficient to merely specify the walls.

An important feature of the programs is the ability to temporarily handle regions where no sensible estimation of parameters can be made. For example, examination of eq 5 and 7 indicates that the estimates will be undefined if the parabola parameter  $c$  is not negative. Also, the quadratic fit is meaningless if less than three data points are regressed. (Fewer data points may be used in some iterations than are defined by the walls; corrected absorbances that are negative are discarded since their  $\log_e$  is not defined.) In these cases



**Figure 1.** First four iterations of the algorithm on a 3-band problem without area renormalization (top) and with area renormalization (bottom). The dotted line is the synthetic spectrum.

the parameter estimates for the band are not updated. Although errors in the estimates—or lack of estimates altogether—may prevent the parameterization of a band during an iteration, the band may be subsequently parameterized if the estimates for other bands improve.

The programs produce two different types of measures of goodness of fit. Although these measures are not employed by the algorithm, they aid the user in determining when sufficient accuracy has been obtained and in judging the suitability of the model. One measure is the regression coefficients from the least-squares fit to the transformed walled data sets. The regression coefficients probably cannot be rigorously interpreted in this application, but experience has shown that low regression coefficients indicate either inappropriateness of the Gaussian model, hidden bands, or lack of convergence; see Table II.

The other measures are based on the root mean squared error (rmse) between the fitted spectrum and the experimental spectrum. This is defined as

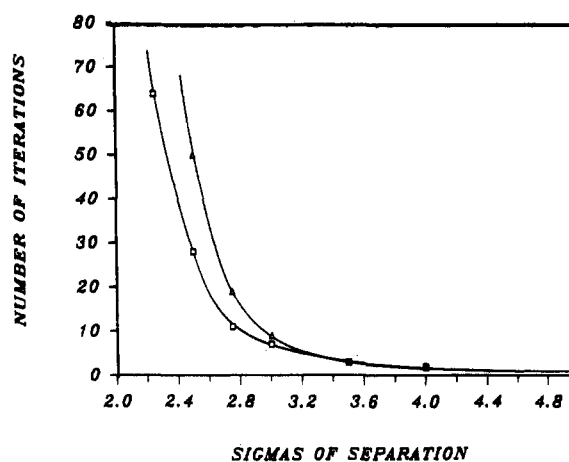
$$\text{rmse} = ((E_i - O_i)^2/n)^{1/2} \quad (8)$$

where  $E_i$  is the height of the fitted spectrum at point  $i$ ,  $O_i$  is the height of the observed spectrum at point  $i$ , and  $n$  is the number of data points included in the sum. The rmse is similar to the objective functions commonly minimized by other resolution methods. We prefer the rmse to other measures for three reasons: it is easily interpreted with respect to the original spectrum; different rmse's may be directly compared independent of the number of data points; and it simplifies the noise analysis.

A partial rmse is calculated for each region defined by a wall set. Each partial rmse gives some indication of the goodness of fit in the region of each proposed band. A total rmse across the entire spectrum is also provided. Partial rmse's that differ significantly from the total rmse are a sign of an incorrect number of proposed bands, failure in the assumption of Gaussian contour, or lack of convergence. Since the time required to compute the partial and total rmse's is often significantly longer than the time required to perform parameter estimation, their calculation after each iteration is optional.

## RESULTS

Figure 1 presents the resolution of a synthetic spectrum containing three overlapped bands with and without area renormalization. The synthetic spectrum contains three Gaussian bands with the same standard deviations centered 3.0 standard deviations apart and with heights of 1.0, 0.5, and 0.25. Ringing of the fit obtained without area renormalization



**Figure 2.** Number of iterations required to achieve 1% accuracy for all parameters with area renormalization (squares) and without area renormalization (triangles) as a function of the separation of the bands. For reference, the data points at 3.0 $\sigma$  of separation correspond to the resolutions presented in Figure 1.

is evident. The ringing is greatly reduced by the use of area renormalization. While both analyses eventually converge to the same result, the resolution performed with area renormalization is considerably faster. Figure 2 presents the rate of convergence for synthetic spectra similar to that of Figure 1 as a function of the separation of the bands and again with and without area renormalization. The data in this graph are naturally quite dependent upon the synthetic spectra being resolved, but we have found that the trends indicated do reflect the typical behavior of the algorithm. In particular, area renormalization usually allows the resolution of more difficult spectra and almost always speeds convergence. Note that area renormalization does not add significantly to the time required to perform an iteration. The numerical integration is not computationally expensive and need only be performed once, since it will not change in subsequent iterations unless the placement of the walls is altered.

Figure 3 presents the resolution of a complex 16-band synthetic spectrum similar to the one used by Pithia and Jones (2). The precision of the absorbances has been limited to  $\pm 0.0001$  by truncation (see Table I for parameters). The top graph shows the underlying bands and the spectrum envelope. The bottom graph shows the accuracy of the fit as a function of time and iteration. Resolution to the limit of the precision of the data is achieved in under 19 min. No initial parameterization other than the placement of walls was given. The times listed in the figures are based upon our measurements

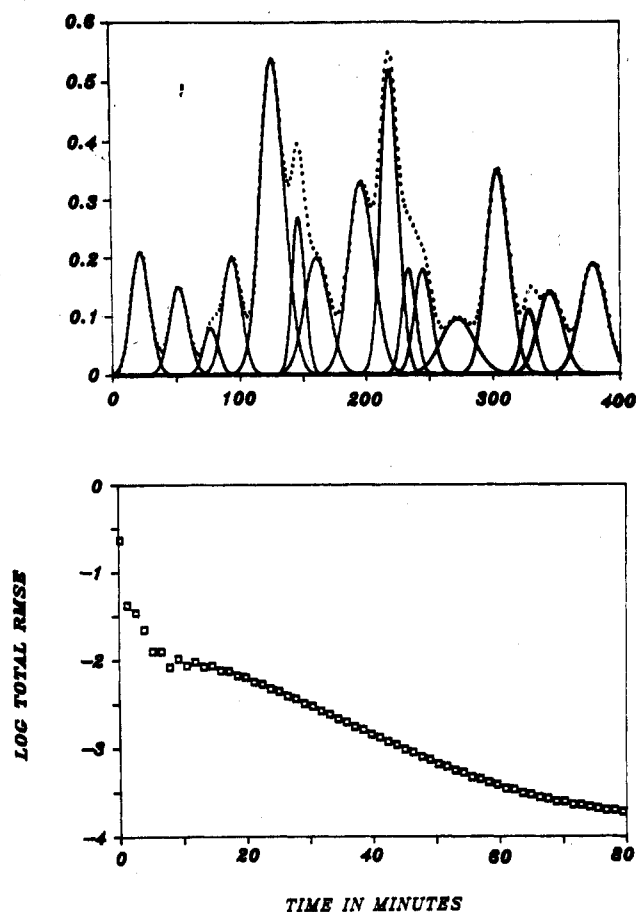


Figure 3. Synthetic test spectrum containing 16 overlapped bands (top). The dotted curve is the spectrum envelope. Time for convergence (bottom). Each square is an iteration.

Table I. Synthetic Test Spectra Parameters for Figure 3<sup>a</sup>

band	mean	std dev	height	area	walls	
1	22.0	7.1	0.210	3.7373	19.0	31.0
2	51.8	7.1	0.150	2.6695	46.0	58.0
3	76.8	5.7	0.080	1.1430	70.0	82.0
4	93.6	7.2	0.200	3.6095	88.0	100.0
5	126.4	9.5	0.540	12.8590	118.0	130.0
6	146.6	5.0	0.270	3.3839	142.0	154.0
7	160.6	9.7	0.200	4.8628	160.0	172.0
8	196.0	10.0	0.330	8.2718	190.0	202.0
9	218.8	6.8	0.520	8.8634	214.0	226.0
10	233.6	5.5	0.180	2.4815	229.0	241.0
11	244.8	6.7	0.180	3.0229	242.0	254.0
12	272.0	13.0	0.095	3.0956	268.0	280.0
13	304.0	8.7	0.350	7.6326	298.0	310.0
14	328.2	5.5	0.110	1.5165	323.0	335.0
15	344.8	9.4	0.140	3.2987	341.0	353.0
16	378.8	10.0	0.190	4.7625	367.0	383.0

<sup>a</sup>Sixteen-band test spectrum.

of the algorithm running on an IBM PC and may differ in implementations on other computers or in other languages. The time required to compute the total rmse's is not included because these are usually omitted particularly in early iterations. The partial rmse's computation time is included and accounts for about half of the duration of each iteration.

Figure 4 shows a base-line-corrected 15-band region of a physiologic fluid amino acid chromatogram. The analysis was performed with a Spectra Physics 8700 solvent delivery system on a Pickering Laboratories 0373150 divinylbenzene cross-linked polystyrene cation exchange column. Detection was by postcolumn development with *o*-phthalaldehyde followed by monitoring fluorescent emission at 430–470 nm with a

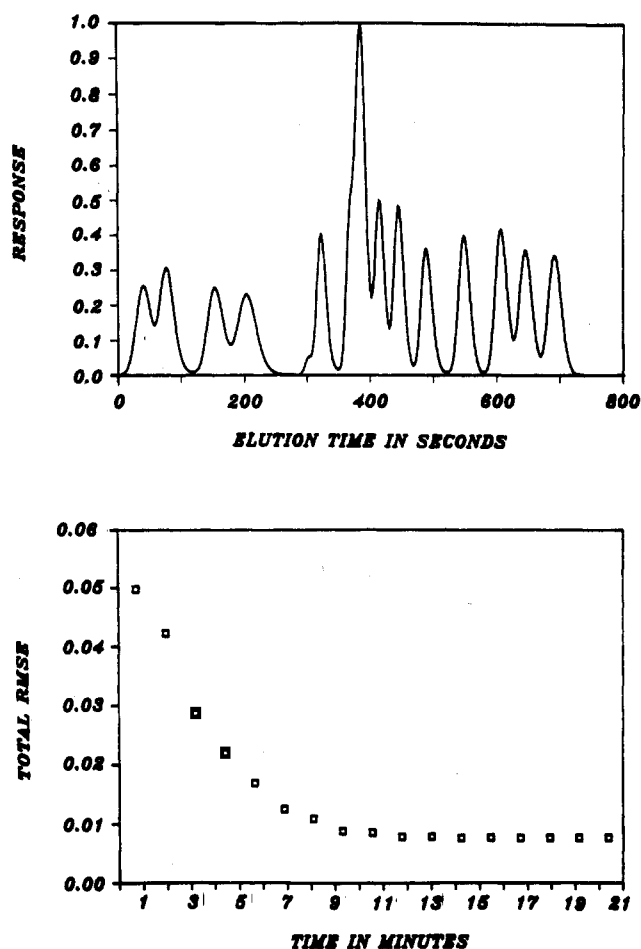


Figure 4. Fifteen-band region of an ion exchange physiologic fluid amino acid chromatogram detected by fluorescence at 430–470 nm after postcolumn development (top). Time for convergence (bottom). Each square is an iteration.

Gilson 121 fluorometer. The gradient program was modified to intentionally reduce the normally excellent separation produced by this method. Figure 4 also shows the total rmse obtained with the new algorithm as a function of time and iteration. The walls were chosen by examination of the chromatogram and were not later modified to attempt to yield better fit. No initial parameterization other than the walls was specified.

Figure 5 shows the difference between the original amino acid chromatogram and the result obtained from the 15th iteration with the new algorithm (see Table II for parameters and walls). The same type of error plot is also shown for the result obtained from a simplified direct descent algorithm using the final parameters from the new algorithm as a starting point.

## DISCUSSION

This new algorithm requires, in the worst case when the wall sets together cover the whole spectrum, roughly the equivalent of one total rmse calculation per iteration to parameterize all bands. Gradient methods require  $3K$  evaluations of a total rmse like calculation per iteration for  $K$  bands in a spectrum. This large increase in computation per iteration is somewhat offset by the relatively few iterations required by most gradient methods when the initial parameterization is fairly good. Simplex methods require at least one total rmse like calculation per iteration, but often require a very large number of iterations. In tests with the new algorithm, convergence to good solutions was obtained on complex spectra usually in under 100 iterations and often less than 20. Therefore the new algorithm should be considerably

Table II. Parameterizations of Spectrum of Figure 4

15th Iteration Parameterization <sup>a</sup>							
band	mean	std dev	height	partial root mean square	RSQRD <sup>c</sup>	walls	
1	40.8	11.23	0.255	0.0016	0.9996	21.0	55.0
2	77.8	11.66	0.299	0.0065	0.9941	61.0	99.0
3	154.3	12.17	0.250	0.0030	0.9990	131.0	175.0
4	204.5	15.35	0.226	0.0031	0.9975	181.0	233.0
5	301.7	5.50	0.047	0.0002	0.9999	293.0	305.0
6	324.1	7.59	0.396	0.0139	0.9906	309.0	339.0
7	367.5	5.83	0.326	0.0076	0.9965	355.0	373.0
8	385.0	8.61	0.973	0.0137	0.9987	377.0	401.0
9	415.7	7.36	0.498	0.0027	0.9998	405.0	429.0
10	446.2	8.11	0.468	0.0116	0.9954	433.0	463.0
11	490.0	8.81	0.350	0.0152	0.9871	473.0	509.0
12	549.4	8.92	0.396	0.0076	0.9973	531.0	567.0
13	607.6	9.22	0.422	0.0071	0.9980	589.0	625.0
14	646.4	10.15	0.355	0.0036	0.9994	629.0	667.0
15	692.7	10.55	0.341	0.0060	0.9976	673.0	715.0

Direct Minimization Parameterization <sup>b</sup>							
band	mean	std dev	height	band	mean	std dev	height
1	40.7	11.05	0.256	9	415.6	7.48	0.497
2	77.5	11.63	0.302	10	445.9	7.88	0.479
3	154.0	12.23	0.249	11	489.4	8.64	0.359
4	204.4	15.33	0.227	12	549.1	8.93	0.397
5	300.7	5.25	0.047	13	607.3	9.28	0.419
6	323.7	7.60	0.398	14	646.2	10.08	0.357
7	367.4	5.54	0.346	15	692.4	10.48	0.344
8	384.9	8.37	0.988				

<sup>a</sup>Total rmse = 0.0076, maximum absolute error = 0.0278. <sup>b</sup>Total rmse = 0.0057, maximum absolute error = 0.0179. <sup>c</sup>Regression coefficient squared.

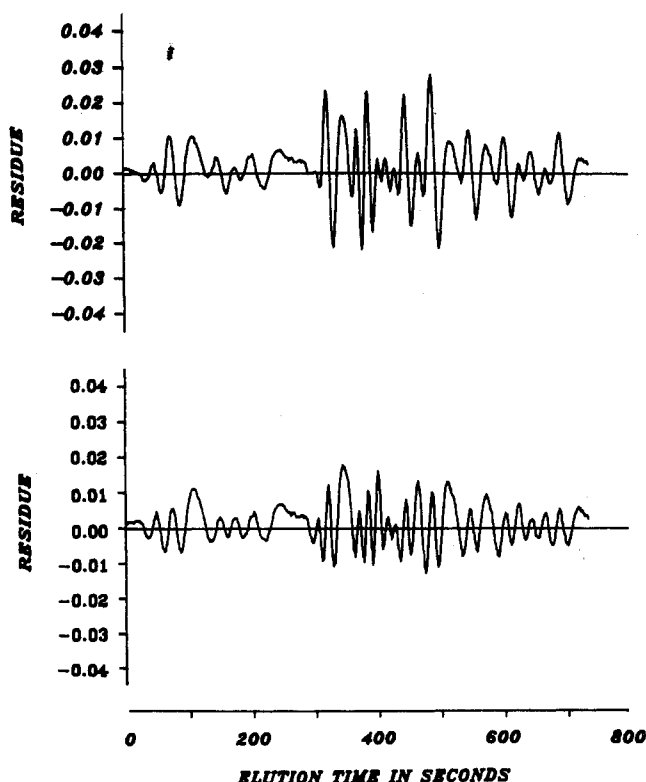


Figure 5. Residue obtained from the 15th iteration parameterization with the new algorithm (top). Residue obtained after additional improvement in the parameterization by a direct minimization routine (bottom).

faster than either gradient or simplex methods. A comparison with the analysis times required by the fastest algorithms tested by Pithia and Jones (2) indicates that the new algorithm is 5–100 times faster. This speed improvement, combined with

modest memory requirements, makes the new algorithm practical for microcomputers.

Model failure places a lower bound on the quality of fit that the new algorithm can obtain. The algorithm relies upon the band shape when finding the best parameterization instead of searching explicitly for an error space minimum. As seen in Figure 5, the gradient method has yielded a small improvement in fit as measured by total rmse. The error space minimization methods will usually yield better fit where there is appreciable model failure. Significant model failure, however, makes the usefulness of any parameterization questionable for most applications.

Area renormalization speeds convergence and enables the resolution of more difficult spectra. It can, however, lead to small errors in the parameter estimates when compared to the estimates that would be obtained without area renormalization because of the inaccuracy of the numerical integration. The magnitude of this error is dependent upon the data interval and the widths of the bands. We have found that this error is usually insignificant and that where there is model failure area renormalization actually improves fit. An option has been provided in programs to disable area renormalization if desired.

The performance of the new algorithm is dependent upon wall placement. It should be clear that the walls are, in a sense, an initial parameterization of the spectrum. We have found that convergence is most likely and most rapid when walls select the relatively narrow regions where the bands are least perturbed by other bands and omit base line. For spectra where there are clear minimas between bands reasonable wall positions are at the minima. For more overlapped spectra the placement of walls is a more difficult and critical process. Most iterative resolution methods are, however, very sensitive to the quality of the initial estimates, this sensitivity increasing dramatically with the difficulty of the problem. Although it becomes more difficult to place walls as the overlap of the

spectrum increases, it is probably easier than estimating all of the parameters of each of the bands as many algorithms require.

It is worth noting that both area renormalization and walls should be applicable to other methods of resolving spectra, including the gradient techniques. Area renormalization could serve to add additional constraints to the parameter estimates, and the judicious placement of walls with the error space minimization methods should make convergence more likely and quicker. The final iterations could still be performed without the walls to ensure that best estimates are obtained. Some references to walls used for similar purposes are present in the literature (9, 11).

We are currently working to adapt this method to other band types. An implementation that is able to resolve Lorentzian as well as Gaussian bands has already been developed. We are also attempting to better characterize the accuracy of this method in the presence of noise and model failure. The results of this work will be presented in a future paper.

#### ACKNOWLEDGMENT

We thank Dennis L. Tebbe for early conversations that helped crystallize some of the ideas presented here, Pickering

Laboratories of Mountain View, CA, for their generous assistance and patience in the acquisition of the chromatogram presented in Figure 4, and Spectra-Physics of San Jose, CA, for aid in interfacing their integrator to our computer.

#### LITERATURE CITED

- (1) Allen, Geoffrey, C.; McMeeking, Robert F. *Anal. Chim. Acta* **1978**, *103*, 73-108.
- (2) Pithia, J.; Jones, Norman R. *Can. J. Chem.* **1966**, *44*, 3031-3050.
- (3) Fraser, R. D. B.; Suzuki, Eikichi *Anal. Chem.* **1966**, *38*, 1770-1773.
- (4) Bartecki, Adam; Soltowski, Janusz; Kurzak, Krzysztof *Comput. Enhanced Spectrosc.* **1983**, *1*, 31-38.
- (5) Lam, C. F.; Forst, Art; Bank, Harvey *Appl. Spectrosc.* **1979**, *33*, 273-278.
- (6) Binsch, Gerhard *Comput. Methods Chem. Proc. Int. Symp.* **1980**, 15-36.
- (7) Slavic, Ilfan A. *Nucl. Instrum. Methods* **1976**, *134*, 285-289.
- (8) Abramowitz, Stegun "Handbook of Mathematical Functions"; National Bureau of Standards: Washington, DC, 1970; p 932.
- (9) Roberts, S. M.; Wilkinson, D. H.; Walker L. R. *Anal. Chem.* **1970**, *42*, 886-893.
- (10) Barker, B. E.; Fox, M. F.; Hayon, E.; Ross, E. W. *Anal. Chem.* **1974**, *46*, 1785-1789.
- (11) Anderson, Andrew H.; Gibb, Terence C.; Littlewood, Anthony B. *J. Chromatogr. Sci.* **1970**, *8*, 640-646.

RECEIVED for review July 23, 1984. Resubmitted January 9, 1986. Accepted January 9, 1986.

## Error Propagation and Figures of Merit for Quantification by Solving Matrix Equations

Avraham Lorber

Nuclear Research Centre—Negev, P.O. Box 9001, Beer-Sheva 84190, Israel

**Quantitation from one-dimensional data enables the simultaneous determination of all components contributing to the spectrum. However, the applicability of the procedure is limited because of the figures of merit; error propagation, signal to noise, limit of detection, precision, accuracy, sensitivity, and selectivity are not determined for each component. It is suggested that by considering the "net analyte signal", error propagation and other figures of merit are defined for each component. Net analyte signal is defined as the part of the signal that is orthogonal to the spectra of the other components. The mathematical results were applied to absorbance data of a four RNA nucleotides mixture, and it was found that it succeeds well in predicting both precision and accuracy.**

Determination of a multicomponent mixture by using the full spectrum of the sample can be accomplished by using linear algebra methods. The benefits of using several data points are resolution of overlapping spectrum (1) and accurate background subtraction (2). However, until now, it was impossible to estimate both the precision and accuracy in determining each component when matrix manipulations are involved. This shortcoming imposes a severe restriction on the applicability of matrix computation methods for quantitative analysis. One can hardly speak about quantitative analysis without knowledge of the amount of signal, error, and related quantities for each component.

The signal-to-noise ratio (S/N) of an analytical system is the most useful figure of merit that can be used to characterize

an analytical technique for a specific application. The S/N is related to three other features of merit: (a) the precision expressed as percent relative standard deviation (RSD) of the concentration measurement; (b) the limit of detection (LOD), which is the concentration of analyte that corresponds to a  $S/N = 3$  (3); and the sensitivity of the analytical method, which is the instrumental response to a certain concentration of analyte and is measured as the slope of the analytical calibration curve. Quantitation from an overlapped spectrum necessitates the definition of two additional figures of merit: the error propagation, which is the ratio of the precision in the determined concentration to the precision of the instrumental response; and the selectivity (4-6), which measures the possibility of deconvoluting the overlapped spectrum.

In fact, a boundary for the error in solving matrix equations is given by the condition number,  $\kappa$ .  $\kappa$  is usually computed as the ratio of the largest to the smallest eigenvalues of the matrix. The product of  $\kappa$  by the error in the measured values gives an upper bound on the error in the determined concentrations vector. This approach of estimating error propagation was introduced to analytical chemistry problems by Jochum et al. (7). The condition number tool was introduced in numerical analysis to estimate the possibility of inverting a matrix and usually results in over estimation by up to 1 order of magnitude (8). Thus, this approach may be regarded as a qualitative tool for error estimation.

A more severe restriction that prevents its practical application to analytical quantification is that the estimate is for the sum of the errors. To demonstrate this problem, we consider a three-component mixture. The spectra of two of