Computational Methods for Resolution of Mass Spectra.

2 AUTHORS, INCLUDING:

David Luenberger

Stanford University

131 PUBLICATIONS   **15,600** CITATIONS

Computational Methods for Resolution of Mass Spectra

DAVID G. LUENBERGER

Stanford Electronics Laboratories, Stanford, Calif.

ULRIC E. DENNIS

General Electric Co., San Jose, Calif.

▶ Computer aided mathematical techniques for the improvement of mass spectrometer data resolution are presented and discussed. Assuming that an output spectrum consists of a linear superposition of pulses, each characteristic of an ion's mass to charge ratio and relative abundance, a "best" relative abundance for each ion in a sample is calculated using a minimum mean square error criterion. Two computational methods, those of normal equations and of quadratic programming, are presented together with an example that compares their effectiveness. The characteristic pulse shapes are shown to be almost Gaussian and based on such an assumption the ability of these methods to resolve two adjacent mass peaks in the presence of noise is discussed. It is concluded that for less than 40% overlap the effect of noise on the ability to resolve is small, while for overlap greater than about 70% resolution should be extremely difficult.

Figure 1.   Time-of-flight mass spectrometer

I N THE ANALYSIS of chemical substances by mass spectrometric techniques a time-of-flight mass spectrometer is often employed because of the advantages which it offers with respect to size, weight, and speed of analysis. A serious disadvantage, however, is its inherently poor resolution. To overcome this disadvantage, certain improvements in the spectrometer's ion gun have been made (5) and, associated analog devices have been constructed (2) to improve resolution. With the increased availability of high speed digital computers it has now become attractive to use computational techniques to improve resolution, as will be demonstrated in this paper.

For a sample containing ions of various mass-to-charge ratios, the ideal spectrum would consist of a series of discrete, narrow pulses, one for each mass-to-charge ratio present. The amplitude of each pulse would equal the relative abundance of its corresponding ion and each pulse's position on the time axis would be a function of the ion's mass-to-charge ratio. Because of the physical limitations of the system, however, these pulses tend not to

be narrow but broad. They may, in fact, become so broad that pulses sufficiently close together overlap, thus giving rise to a continuous rather than a discrete spectrum. This spectrum may, in turn, be so smooth that the determination by visual inspection of relative abundances or even of the mass peaks themselves becomes extremely difficult.

It will be assumed that for a sample consisting only of ions of a single mass-to-charge ratio the spectrum will be a pulse whose shape is a characteristic of the spectrometer and its control settings. For a sample containing ions of various mass-to-charge ratios the spectrum would be a superposition of such characteristic pulses. The area under each pulse would be equal to the corresponding ion's relative abundance and the time of occurrence of each pulse peak would be a function of the ion's mass-to-charge ratio. Overlapping will occur for pulses sufficiently close together.

If one knows the characteristic output pulse of the spectrometer, then by some relatively simple computational procedures the spectrum may be resolved and the individual mass peaks obtained. One complication, however, is the inherent noise in the system. Meter, chart, and oscilloscope readings are not absolutely precise. There is noise in the amplifiers; power supply voltages may fluctuate; the vacuum systems aren't perfect, etc. If for a particular sample it is found that up to n distinct mass peaks may exist one might ordinarily set up n simultaneous equations to solve for the relative abundances. However, due to the presence of noise this may lead to erroneous results. Therefore, it be-

comes advantageous to set up more equations than there are unknown mass peaks and minimize the effects of this noise by computing the best estimates to the relative abundances in a mean square sense.

It is the purpose of this paper to briefly examine the shape of the characteristic pulse, introduce two methods for resolving the mass spectrum based on the linear superposition model of the spectrometer, and analyze the resulting errors.

TIME-OF-FLIGHT MASS SPECTROMETER

Although the basic operation of the time-of-flight mass spectrometer has been well described in the literature, it will be reviewed here briefly so that the derivation of the characteristic pulse shape may be better understood. A more complete description can be found in (6).

Referring to Figure 1, the sample to be analyzed is introduced into the ionization region where an electron beam causes some of the sample to become ionized. The initial positions of these ions are randomly distributed with this region, a given ion being located s centimeters from the end of the region as shown. A voltage pulse on one of the grids creates an electric field E_s which causes the positively charged ions to move into the acceleration region. There, a high electric field E_d causes the ions to accelerate so that they pass into the long, field free drift region. Ions with different mass-to-charge ratios will have attained different velocities when entering the drift region and thus will arrive at the ion collector at different times. The
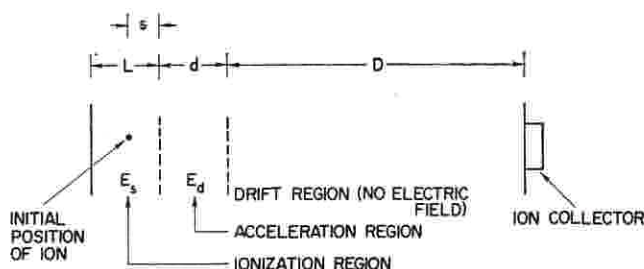
Table I. Typical Mass Spectrometer Parameters

Parameter	Symbol	Typical value
Length of drift region	D	40 cm.
Length of acceleration region	d	1.2 cm.
Length of ionization region	L	0.4 cm.
Electric field, ionization region	E_s	320 volts/cm.
Electric field, acceleration region	E_d	1280 volts/cm.

greater the mass-to-charge ratio the longer will be the drift time. When the resulting current at the ion collector is plotted against time, the mass spectrum is obtained.

Typical parameters as reported by Wiley and McLaren (6) are given in Table I.

CHARACTERISTIC PULSE SHAPE

Ideally, the output spectrum of a time-of-flight mass spectrometer should be a sequence of sharp, narrow pulses, each pulse corresponding to a different mass-to-charge ratio and having an amplitude proportional to its relative abundance in the sample. For several reasons, primarily variations in the initial position and initial velocity of the ions along the flight axis, these sharp pulses tend to broaden. Given a collection of ions having mass-to-charge ratios sufficiently close together their corresponding pulses may overlap to such a degree that they may not be clearly distinguished by the unaided eye.

Using the notation of Wiley and McLaren (6) the time of flight t, is given by

$$t = t_s + t_d + t_D \quad (1)$$

where t_s, t_d and t_D are, respectively, the times of flight in the ionization, acceleration, and drift regions. Using Newton's second law of motion, it may be shown that

$$t_s = 1.02 \frac{(2m)^{1/2}}{qE_s} [(u_o + qsE_s)^{1/2} \pm (u_o)^{1/2}] \quad (2)$$

$$t_d = 1.02 \frac{(2m)^{1/2}}{qE_d} [(u)^{1/2} - (u_o + qsE_s)^{1/2}] \quad (3)$$

$$t_D = 1.02(2m)^{1/2} \frac{D}{2(u)^{1/2}} \quad (4)$$

Here u_o is a kinetic energy computed from the component of the initial velocity along the axis of flight and

$$u = u_o + qsE_s + qdE_d \quad (5)$$

The choice of plus or minus in Equation 2 depends on whether the initial axial velocity is opposed to or in agreement with the ultimate direction of flight. The factor 1.02 results from expressing time in microseconds, m, the ion's mass, as a mass number, u and u_o in electron volts and q, the ion's charge, in number of electronic charges.

The pulse shape caused by variations in initial energy will first be examined. It will be assumed that all ions have the same initial position. Assuming a Maxwell-Boltzman distribution in the initial velocity of the ions (6) and letting v_o be the component of velocity along the axis of flight the number of ions with initial axial velocity between v_o and $v_o + dv_o$ is given by (3)

$$dNv_o = \frac{N}{\sqrt{\pi}} \left(\frac{m}{2kT} \right)^{1/2} exp \left(-\frac{mv_o^2}{2kT} \right) dv_o \quad (6)$$

where

N = total number of ions present
m = mass of ion
k = Boltzman's constant
T = absolute temperature

In terms of initial energy this becomes

$$dNu_o = \frac{N}{\sqrt{\pi}} \left(\frac{1}{4u_o kT} \right)^{1/2} exp \left(-\frac{u_o}{kT} \right) du_o \quad (7)$$

From the expression for the time of flight and for the typical parameters given previously, it can be shown that approximately

$$\frac{du_o}{dt} = \frac{445}{\sqrt{m}} \sqrt{|u_o|} \quad (8)$$

assuming singly charged ions and that $0.1 \leq s \leq 0.3$. From Equations 1, 7, and 8, it may be shown that approximately

$$i = \frac{6.04 \times 108\, Ne}{\sqrt{m}} exp \left[- \left(1068 \frac{t}{\sqrt{m}} - 64.5 \sqrt{s} - 818 \right)^2 \right] \quad (9)$$

where

e = charge on the electron in coulombs
m = mass number of ion
i = current in amperes
N = total number of ions
t = time in microseconds
s = initial position of ion in centimeters

Hence, for all ions located at a given initial position s the output will be a pulse of almost Gaussian shape centered at time t_o such that

$$\frac{t_o}{\sqrt{m}} = 0.766 + 0.0604 \sqrt{s} \quad (10)$$

Since all ions do not have the same initial position, there will be some further distortion of the output pulse and this effect must be taken into account when considering Equations 9 and 10. The derivation of an analytical expression in this case becomes somewhat difficult because of discontinuities encountered in some of the analytical expression. However, computations were performed graphically and the results indicated that the characteristic pulse remained essentially Gaussian but with some additional asymmetrical distortion. This conclusion was sufficient to permit an analysis of the errors arising from the computational techniques to be presented.

METHODS OF CALCULATION

Consider the time scale shown in Figure 3(a) and let t_i ($i = 1, 2, \ldots, 5$) represent some arbitrary fixed points on the scale. At each t_i there is centered a pulse of amplitude x_i. A missing pulse is considered to be a pulse of zero amplitude, as for $i = 2$ in Figure 2(a). The shapes of all pulses are known but need not be the same. When they are superimposed—i.e., at every point in time their amplitudes are added together—the spectrum of Figure 2(b) is obtained. In the resolution of a mass spectrogram the reverse procedure is applied. Knowing the mass spectrum, points on the time axis, and pulse shapes the problem is to determine the amplitudes x_i. A knowledge of the x_i permits the calculation of the areas under the pulses which in turn equal the relative abundances. The points t_i are indicative of the mass-to-charge ratios.

In the discussion to follow it will be assumed (without loss of generality) that all ions from a sample are singly ionized. Hence, one may consider a sample as containing ions of various mass numbers rather than mass-to-charge ratios. Assume that in the section of spectrum to be analyzed there may be n mass numbers present. Let x_i be the amplitude of the ith mass number's characteristic pulse, $i = 1, 2, \ldots, n$. Note that $x_i \geq 0$, the equal sign applying when ions of a particular mass number are not present.

Consider the spectrum shown in Figure 3 and let it be divided into m discrete points. At the jth point, the amplitude, y_j, is equal to a sum of contributions made by each of the n mass numbers in the spectrum. Each contribution is proportional to the relative abundance of the contributing mass number's ions. Let the contribution of the ith mass number to the amplitude of the spectrum of the jth point be $a_{ji}x_i$. Then, in general, one can write
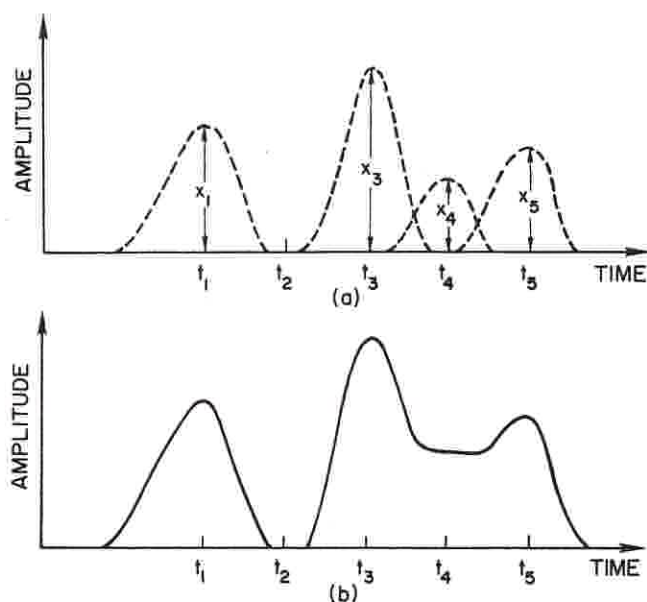
Figure 2. Mass spectrum
(a) Superposition of pulses
(b) Complete spectrum

Figure 3. Observed mass spectrum

$$y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n$$
$$y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n$$

$$\cdot$$
$$\cdot \qquad (11)$$
$$\cdot$$

$$y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n$$

In matrix notation this becomes

$$y = Ax \qquad (12)$$

where y is an $m \times 1$ column vector, x is an $n \times 1$ column vector, and A is an $m \times n$ matrix. The elements of A are calculated directly from a knowledge of the characteristic pulses, the locations of possible mass peaks, and the spacing between the y_i.

If $m = n$, the number of equations would equal the number of unknowns and the solution would simply be $x = A^{-1}y$. But because of noise, the solution would be unreliable and since more information than just n points on the spectrum are known, there should be more accurate methods of computing the amplitudes. Hence, one may set up more equations than there are unknowns ($m > n$) and solve them in some optimal way. Two such methods are indicated below.

The criterion for the optimal solution is as follows. It can be assumed that the noise in the system has zero mean, that is, the readings y_i are just as likely to be too high as they are too low. With Equation 12 corresponding to $\bar{x}$ an estimate to the solution, one can calculate $\bar{y}$, or what the readings should be if there were no noise. The difference between y, the noisy observation, and $\bar{y}$, the noise free estimate to what the readings should be, is an error vector e. It will be assumed that the best estimate to

the solution, $\hat{x}$ is the one that makes $\|e\|^2$ a minimum, where $\|e\|$ is the norm of the error. The norm of an n-dimensional vector e, whose components are e_i ($i = 1, 2, \ldots, n$), is given by

$$\|e\| = \left(\sum_{i=1}^{n} e_i^2 \right)^{1/2} \text{ and may be considered the magnitude of the vector in the } n\text{-dimensional space.}$$

In the examples to follow, two assumptions were made to simplify the computation of the characteristic matrix A. If the mass numbers involved are not too small and the range of mass numbers over which the analysis is to be carried out is not too great, then the characteristic pulse shape may be considered constant over this range and consecutive mass peaks on the time scale equally spaced apart. In addition to simplifying the computation of A, these assumptions permit the relative abundances to be given by the pulse amplitudes rather than just the pulse areas. If greater accuracy is desired or if analysis over the same range of mass numbers is frequently carried out, it may become desirable to make a more accurate computation of the matrix A.

METHOD OF NORMAL EQUATIONS

The matrix equation $y = Ax$ consists of m linear equations in n unknowns where $m > n$. Because of noise in the system, these equations are not all consistent. That is, if one took two groups of n equations each and solved them simultaneously, he would obtain two different sets of solutions. Since without any other a priori knowledge, it must be assumed that
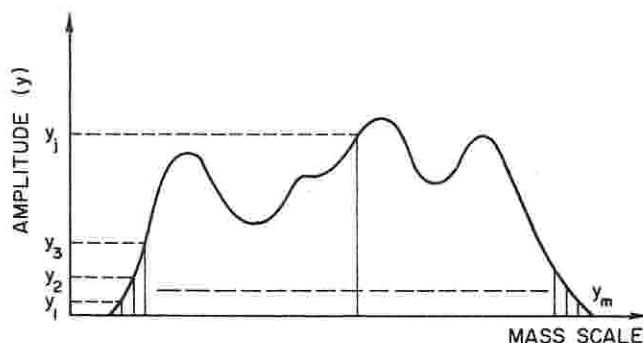
each of the m equations are equally valid, the best that can be done is to choose an $x = \hat{x}$ in such a way that $\hat{x}$ is an optimum solution in some sense. One criterion for an optimum solution is as follows. If in each of the m equations $y_i = \sum_j a_{ij}x_j$ the optimum solution $\hat{x}$ were substituted, then one obtains $\hat{y}_i = \sum_j a_{ij}x_j$ where, in general, $\hat{y}_i \neq y_i$. Hence, there is an error e_i whose square is

$$e_i^2 = (y_i - \hat{y}_i)^2 \qquad (13)$$

If the squared errors for the m equations are added together, one obtains

$$\|e\|^2 = \sum_{i=1}^{m} e_i^2 \text{ or, in matrix notation}$$

$$\|e\|^2 = (y - A\hat{x})^T(y - A\hat{x}) \qquad (14)$$

where the superscript T indicates the transpose of a matrix. The optimum solution x is the one that will make $\|e\|^2$ a minimum. The solution to this problem is (see Appendix I)

$$\hat{x} = (A^TA)^{-1}A^Ty \qquad (15)$$

Although these calculations are straightforward, there is nothing in the method of solution to prevent one or more of the computed relative abundances to come out negative, as illustrated in Table II. To overcome this drawback, the method of quadratic programming may be used to solve $y = Ax$.

Table II. Computed Relative Abundances

Relative abundances
($^{84}Kr = 100$)

Mass number	Computed by		Handbook value
	Normal eqs.	Quad. prog.	
78	1.2	1.2	0.6
79	0.5	0.5	0
80	4.2	4.1	3.5
81	−0.1	0	0
82	14.1	14.1	20.2
83	16.7	16.9	20.2
84	100.0	100.0	100.0
85	−0.3	0	0
86	28.5	28.4	30.6
87	0.0+	0.1	0
88	0.0−	0	0

It is again desired to choose an x such that the squared error, $\|e\|^2$ is minimized. Hence,

$$\|e(x)\|^2 = (y - Ax)^T(y - Ax) =$$
$$x^T(A^T A)x - 2y^T Ax + y^T y \quad (16)$$

must be minimized with respect to x, subject to the constraint $x \geq 0$ (no negative abundances). The method of solution is an iterative procedure whereby a sequence of vectors x^1, x^2, ... are calculated such that they converge to the optimal solution. An initial solution x^0 (≥ 0) is chosen. From this a new solution x^1 is calculated by first computing x_1^1, then x_2^1, then x_3^1, etc., until all n components of x^1 have been determined. In the same manner vector x^2 is computed from x^1 and so on. By noting the rate at which the solution converges, one can terminate the iterative process when the desired accuracy has been obtained.

The procedure for this process, described in (1), is as follows:

$$x_i^{(k+1)} = max(0, w_i^{(k+1)})$$

where

$$w_i^{(k+1)} = -\frac{1}{q_{ii}}\left[\sum_{j=1}^{i-1} q_{ij}x_j^{(k+1)} + \sum_{j=i+1}^{n} q_{ij}x_j^{(k)} - b_i\right] \quad (17)$$
$$Q = A^T A \qquad b = y^T A$$

That is, always calculate $w_i^{(k+1)}$ from the most up-to-date components of x available. For example, in calculating $x_5^{(10)}$, use $x_1^{(10)}$ through $x_4^{(10)}$ and $x_6^{(9)}$ through $x_{12}^{(9)}$. If $w_i^{(k+1)}$ is negative, set $x_i^{(k+1)}$ equal to zero. Otherwise, set $x_i^{(k+1)}$ equal to $w_i^{(k+1)}$. This method is quite simple to implement on a computer and some results are given in Table II.

COMPARISON OF METHODS

The question of which method one should use depends on the given situation. In quadratic programming the optimal estimates of the relative abundances are constrained to be nonnegative and, therefore, if an abundance were actually zero it could never be estimated low. Thus the constraint tends to introduce bias. Consequently, if several runs are to be averaged normal equations rather than quadratic programming should be applied to each run. Quadratic programming could then be applied to the average.

A simple and often effective procedure is to replace negative abundances obtained by normal equations with zero and leave all other values the same. This procedure, however, generally leads to somewhat different answers from those obtained by quadratic programming and therefore gives larger expected squared-errors.
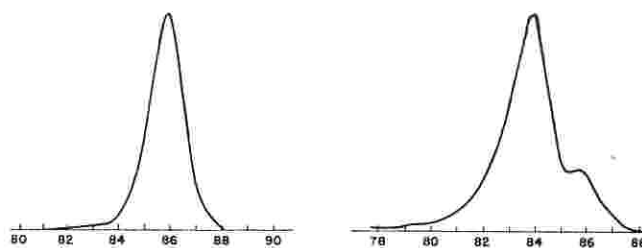

Figure 4. Mass spectra of krytopn

(a) Krypton-84
(b) Normal krypton

EXAMPLES

To illustrate the effectiveness of these two methods of calculation, data were taken from a paper by B. R. F. Kendall (2) and the methods applied. The data were in the form of an output pulse for a sample containing krypton-86 (giving the characteristic pulse) and the spectrum observed for normal krypton. These are shown in Figures 4(a) and 4(b), respectively. Sixty-four equations were solved in eleven unknowns. The results are shown in Table II. The computations were made on the Burroughs B-5500 computer at Stanford University. It may be seen that the small negative relative abundances calculated by the method of normal equations turned out to be zero under quadratic programming. Otherwise, the two methods give results which are in fairly close agreement with each other and with the handbook values.

ERROR ANALYSIS

If there were no noise in the system, then n equations in n unknowns would completely determine a spectrum regardless of how much two adjacent peaks overlapped. Due to the presence of noise, however, this will not be possible and for each measurement made (point in the spectrum) there will be some error. The question now is how broad the characteristic pulses can be (or, to what degree two adjacent pulses may overlap) before the noise limits the ability of the analysis to resolve mass peaks.

Because of noise, every observation y_j is equal to a linear combination of the relative abundances plus some noise,

or, $y_j = \sum_{i=1}^{n} a_{ji}x_i + w_j$ where y_j is the jth element of the y vector, a_{ji} are elements of the A matrix, and w_j is the observation noise of the jth measurement. The complete matrix equation is

$$y = Ax + w \quad (19)$$

where y, A, and x are as before, and w is an $m \times 1$ noise vector. The optimal estimate without the consideration of noise—i.e., $w = 0$—is

$$\hat{x} = (A^T A)^{-1}A^T y \quad (20)$$

The error, defined as $\hat{x} - x$, becomes

$$\Delta x = \hat{x} - x =$$
$$(A^T A)^{-1}A^T[y - (y - w)] =$$
$$(A^T A)^{-1} A^T w \quad (21)$$

Statistically, the error Δx may be expressed by its covariance matrix $E(\Delta x \Delta x^T)$, an $n \times n$ matrix whose elements are $E(\Delta x_i \Delta x_j)$, the expected values of $\Delta x_i \Delta x_j$. This matrix may be expressed as

$$E(\Delta x \Delta x^T) =$$
$$E[(A^T A)^{-1}A^T ww^T A(A^T A)^{-1}] \quad (22)$$

The only random variable on the right side of Equation 22 is w. Hence,

$$E(\Delta x \Delta x^T) =$$
$$(A^T A)^{-1}A^T E(ww^T)A^T A^T A)^{-1} \quad (23)$$

Now, $E(ww^T)$ is an $m \times m$ matrix whose elements are $E(w_i w_j)$. If it is assumed that the measurement errors are uncorrelated with zero mean and that the variance of each measurement is the same ($= \sigma^2$), then

$$E(w_i w_j) = 0 \qquad i \neq j$$
$$E(w_i w_j) = \sigma^2 \qquad i = j \quad (24)$$

and

$$E(ww^T) = \sigma^2 I$$

where I is the identity matrix. Hence,

$$(E\Delta x \Delta x^T) =$$
$$\sigma^2(A^T A)^{-1} A^T A(A^T A)^{-1} =$$
$$\sigma^2(A^T A)^{-1} \quad (25)$$

The covariance matrix of the error is then equal to σ^2 times the inverse matrix used to solve for the optimal estimate itself.

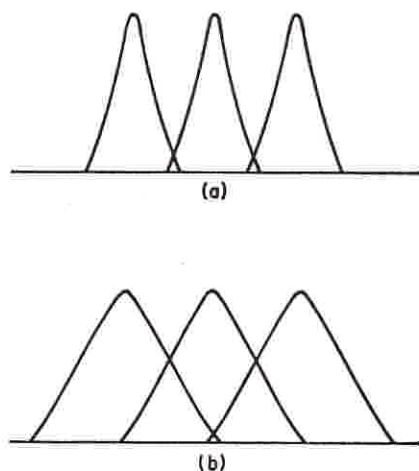For specific values a_{ji}, the elements of A, the matrix $(A^T A)^{-1}$ may be

Figure 5. Correlation between adjacent pulses

(a) Small correlation
(b) Large correlation

easily computed. However, if one assumes an analytical expression for the characteristic pulse shape (such as a Gaussian-shaped pulse) then although the matrix $A^T A$ may be expressed analytically, an analytical expression for its inverse may be hard to find.

Up to this point no simplifying assumptions regarding matrix A have been made and the analysis that leads to Equation 25 has general validity. In fact, the pulse shape need not be Gaussian. Further analysis in such general terms could be carried out at the expense of great complexity and effort. Therefore, to obtain some useful results with reasonable effort, it will again be assumed that over the range of interest each mass-to-charge ratio has the same pulse shape and that the mass peaks are equally spaced along the time axis. The matrix $A^T A$ will then be of the form

$$
A^T A = \begin{bmatrix}
c_0 & c_1 & c_2 & \cdots & c_{n-1} \\
c_1 & c_0 & c_1 & \cdots & c_{n-2} \\
c_2 & c_1 & c_0 & \cdots & c_{n-3} \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
c_{n-1} & c_{n-2} & c_{n-3} & \cdots & c_0
\end{bmatrix} \quad (26)
$$

where the elements c_k are the inner products of the various columns of A and are a measure of how well a given pulse will correlate with another pulse k steps away. Thus, c_0 will be the largest of the c_k and the rest will become successively smaller, or,

$$
c_0 > c_1 > c_2 > \ldots > c_{n-1} \quad (27)
$$

If this sequence is rapidly decreasing there is little correlation or overlap

between pulses, as shown in Figure 5(a). If the sequence decreases slowly, there is greater correlation or overlap, as shown in Figure 5(b). The overlap between two pulses is here defined as follows. If $x(t)$ and $y(t)$ are the expressions for two pulses, as shown in Figure 6, the overlap p is given by

$$
p = \frac{\int_{-\infty}^{\infty} x(t)\, y(t)\, dt}{\left[\int_{-\infty}^{\infty} x^2(t)dt \int_{-\infty}^{\infty} y^2(t)dt \right]^{1/2}} \quad (28)
$$

If the two pulses are identical in shape and separated in time by T seconds Equation 28 reduces to

$$
p = \frac{\int_{-\infty}^{\infty} x(t)x(t-T)dt}{\int_{-\infty}^{\infty} x^2(t)dt} \quad (29)
$$

Note that in both Equations 28 and 29, $|p| \leq 1$. In the discrete case, the overlap is defined in terms of summations rather than integrals. The terms $\frac{c_k}{c_0}$ give the overlap between two pulses k mass numbers apart. The characteristic pulse shape from the mass spectrometer is Gaussian. This leads to a matrix $A^T A$ of the form of Equation 26 with $c_k = p^{2k}$ (where for convenience it is assumed that $c_0 = 1$).

Although a simple analytical expression for $(A^T A)^{-1}$ cannot be found even when it has the known form $c_k = p^{2k}$ it is possible to derive an approximate expression for the inverse valid for very large matrices—i.e., $n \gg 1$. This has been shown in (4).

Of primary interest are the diagonal terms of the matrix $\sigma_0^2 (A^T A)^{-1}$ since they represent the variance of the error in any one estimate. These elements are all the same and their value is denoted by σ_p^2.

It can be shown (4) that as $n \to \infty$,

$$
\sigma_p^2 = \sigma_o^2 \frac{\sum_{k=1}^{\infty} (-1)^{k+1}\, p^{(k-1/2)^2}}{\sum_{k=1}^{\infty} (-1)^{k+1}\, (2k-1)\, p^{(k-1/2)^2}} \quad (30)
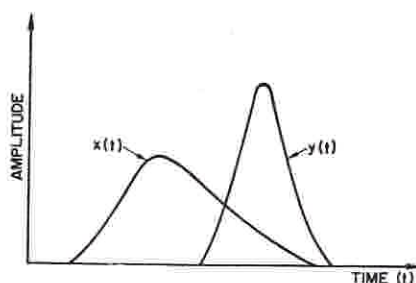$$
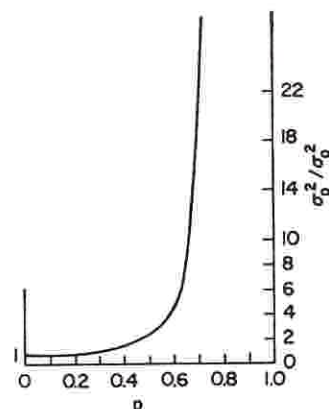

Figure 6. Correlation between two pulses

Figure 7. Noise vs. amount of overlap

This expression, which is plotted in Figure 7 can be used as a fairly reliable estimate of the error in situations where the pulse shape is approximately Gaussian and which involves several possibly overlapping pulses.

The case $p = 0$, corresponds to zero overlap and $\sigma_p^2 = \sigma_o^2$. That is, the variance of the error in the estimate is the same as the variance that would be obtained had each mass number been observed individually by the spectrometer, subject to the same noise conditions. As the overlap between adjacent pulses increases, p increases. For $p \leq 0.4$, $\sigma_p^2 \approx \sigma_o^2$. However, between $p = 0.4$ and $p = 0.6$, the variance will increase by almost a factor of 4. After $p = 0.6$, the variance increases sharply. Hence, there exists a practical limit as to how much overlap can be tolerated until the error in the estimate loses its usefulness. Of course, keeping the measurement noise σ_o^2 low will help somewhat, but due to the sharp increase in σ_p^2/σ_o^2 after $p = 0.6$, this reduction in noise will not be as great a help as one might wish.

The spectrum of normal krypton shown in Figure 4(b) has $p = 0.6$ so reliable estimates are to be expected. If, however, the spacing between mass numbers were reduced by 50%, p would increase to $p \simeq 0.8$ and even the finest measurements would not allow the resolution of the spectrum, since for $p = 0.8$, $\sigma_p^2/\sigma_o^2 \simeq 10^6$.

CONCLUSIONS

The resolution of mass spectrometer data by computational techniques is feasible and may be used to good advantage as long as the overlap between adjacent pulses is not too high. For overlaps greater than about 60%, the error in the estimate may no longer be of value. This will happen regardless of how carefully one might make his measurements.

APPENDIX I

It is given that

$$y = Ax \qquad (11)$$

where y is an $m \times 1$ vector, x is an $n \times 1$ vector, A is an $m \times n$ matrix, and $m > n$. It is desired to find x such that

$$\|e\|^2 = (y - Ax, y - Ax) \qquad (12)$$

is a minimum. Let the vector x which minimizes $\|e\|^2$ be $\hat{x}$. Suppose x is replaced by $\hat{x} + \epsilon \ell$ where ϵ is a parameter and ℓ is an arbitrary $n \times 1$ vector. Then, since $\hat{x}$ is the optimal solution

$$\|e(\epsilon)\|^2 =$$
$$(y - A(\hat{x} + \epsilon \ell),$$
$$y - A(\hat{x} + \epsilon \ell) \geq \|e(0)\|^2 \qquad (13)$$

The squared error $\|e(\epsilon)\|^2$ is a continuous function of ϵ and achieves its minimum at $\epsilon = 0$. Therefore, its derivative with respect to ϵ must be zero at $\epsilon = 0$.

Hence

$$0 = \left. \frac{d\|e(\epsilon)\|^2}{d\epsilon} \right|_{\epsilon = 0} =$$
$$2(A\hat{x}, A\ell) - 2(y, A\ell) \qquad (14)$$

and

$$(A\hat{x} - y, A\ell) =$$
$$(A^T A\hat{x} - A^T y, \ell) = 0 \qquad (15)$$

Since Equation 15 must hold for all ℓ it follows that

$$A^T A\hat{x} - A^T y = 0 \qquad (16)$$

and

$$\hat{x} = (A^T A)^{-1} A^s y \qquad (17)$$

LITERATURE CITED

(1) Hildreth, C., *Nav. Res. Log. Q.* **4** pp. 79–85 (1957).
(2) Kendall, B. R. F., *Rev. Sci. Instr.* **33**, 30 (1962).
(3) Lee, J. F., Sears, F. W., Turcotte, D. L., "Statistical Thermodynamics," p. 53, Addison-Wesley, Reading, Mass., 1963.
(4) Luenberger, D. G., "Resolution of Mass Spectrometer Data," Stanford Electronics Laboratories Report **SU-SEL-64-129**, (Nov. 1964).
(5) McDowell, C. A., "Mass Spectrometry," pp. 29–34, McGraw-Hill, New York, 1963.
(6) Wiley, W. C., McLaren, I. H., *Rev. Sci. Instr.* **26**, 1150 (1965).