Interferences. Methylene chloride or chloroform-soluble carboxylic acids (acetylsalicylic acid, salicylic acid, etc.) which inhibited 4-AAP response were readily extracted from these phases using 0.1*N* sodium hydroxide.

Span-type excipients, which gave a slight response to the 4-AAP reagent, were eliminated by virtue of their insolubility in 10% aqueous sodium chloride; while interfering Tween-type excipients were precipitated (*3*) using the Tween reagent. Propylene glycol, when formulated at a level of 5.25%, contributed ca. 2 to 3% to the observed absorbance. This was minimized by a more favorable distribution of the interference within the aqueous phases used in the procedure.

Placebo analyses indicated that no interference was obtained with such common excipients as stearic acid, stearyl alcohol, cetyl alcohol, petrolatum, methyl or propyl-*p*-hydroxybenzoates, sesame oil, or thimerosal. Lipotropic agents (betaine or choline), all the common vitamins, and neomycin sulfate failed to interfere.

Scope of Reaction. In addition to the other types of steroids reported to react with the 4-AAP reagent, the following steroids reacted at elevated temperature (boiling point of methanol) and at increased concentration (2.0 mg. per 10.0 ml. of reagent): 2α-hydroxymethyl - 17β - hydroxy - 17α - methyl-5α-androst-3-one (306 mμ); and 2α-hydroxymethyl - 17β - hydroxy - 5α-androst-3-one (306 mμ). Steroids without a keto group failed to give any response. The quantitative aspects of the above responses to the 4-AAP reagent were not investigated.

Correlations of Chromophore Wavelength with Structure. The additive effect of various ring *A* and *B* substituents on the chromophore of the parent saturated-3-keto steroid is presented in Table II.

LITERATURE CITED

(1) Cohen, H., Bates, R. W., *J. Am. Pharm. Assoc., Sci. Ed.* **40**, 35 (1951).
(2) Hüttenrauch, R., *Z. Physiol. Chem.* **326**, 166 (1961).
(3) Pitter, P., *Chem. & Ind.* **42**, 1832 (1962).

RECEIVED for review September 4, 1963. Accepted March 13, 1964.

Smoothing and Differentiation of Data by Simplified Least Squares Procedures

ABRAHAM SAVITZKY and MARCEL J. E. GOLAY

The Perkin-Elmer Corp., Norwalk, Conn.

▶ In attempting to analyze, on digital computers, data from basically continuous physical experiments, numerical methods of performing familiar operations must be developed. The operations of differentiation and filtering are especially important both as an end in themselves, and as a prelude to further treatment of the data. Numerical counterparts of analog devices that perform these operations, such as *RC* filters, are often considered. However, the method of least squares may be used without additional computational complexity and with considerable improvement in the information obtained. The least squares calculations may be carried out in the computer by convolution of the data points with properly chosen sets of integers. These sets of integers and their normalizing factors are described and their use is illustrated in spectroscopic applications. The computer programs required are relatively simple. Two examples are presented as subroutines in the FORTRAN language.

THE PRIMARY OUTPUT of any experiment in which quantitative information is to be extracted is information which measures the phenomenon under observation. Superimposed upon and indistinguishable from this information are random errors which, regardless of their source, are characteristically described as noise. Of fundamental importance to the experimenter is the removal of as much of this noise as possible without, at the same time, unduly degrading the underlying information.

In much experimental work, the information may be obtained in the form of a two-column table of numbers, *A vs. B*. Such a table is typically the result of digitizing a spectrum or digitizing other kinds of results obtained during the course of an experiment. If plotted, this table of numbers would give the familiar graphs of %T *vs.* wavelength, pH *vs.* volume of titrant, polarographic current *vs.* applied voltage, NMR or ESR spectrum, or chromatographic elution curve, etc. This paper is concerned with computational methods for the removal of the random noise from such information, and with the simple evaluation of the first few derivatives of the information with respect to the graph abscissa.

The bases for the methods to be discussed have been reported previously, mostly in the mathematical literature (*4*, *6*, *8*, *9*). The objective here is to present specific methods for handling current problems in the processing of such tables of analytical data. The methods apply as well to the desk calculator, or to simple paper and pencil operations for small amounts of data, as they do to the digital computer for large amounts of data, since their major utility is to simplify and speed up the processing of data.

There are two important restrictions on the way in which the points in the table may be obtained. First, the points must be at a fixed, uniform interval in the chosen abscissa. If the independent variable is time, as in chromatography or NMR spectra with linear time sweep, each data point must be obtained at the same time interval from each preceding point. If it is a spectrum, the intervals may be every drum division or every 0.1 wavenumber, etc. Second, the curves formed by graphing the points must be continuous and more or less smooth—as in the various examples listed above.

ALTERNATIVE METHODS

One of the simplest ways to smooth fluctuating data is by a moving average. In this procedure one takes a fixed number of points, adds their ordinates together, and divides by the number of points to obtain the average ordinate at the center abscissa of the group. Next, the point at one end of the group is dropped, the next point at the other end added, and the process is repeated.

Figure 1 illustrates how the moving average might be obtained. While there is a much simpler way to compute the moving average than the particular one described, the following description is correct and can be extended to more sophisticated methods as will be seen shortly. This description is based on the concept of a convolute and of a convolution function. The set of numbers at the right are the data or ordinate values, those at the left, the abscissa information. The outlined

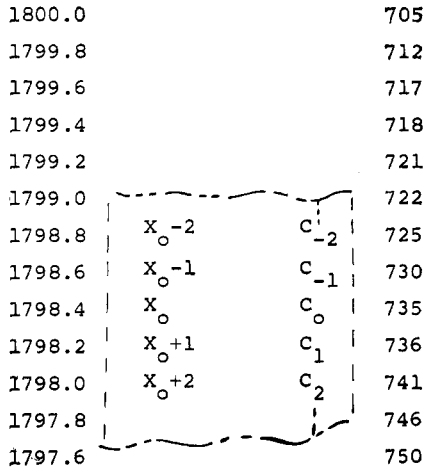

Figure 1. Convolution operation

Abscissa points at left, tabular data at right. In box area the convolution integers, c_i. Operation is the multiplication of the data points by the corresponding C_i, summation of the resulting products, and division by a normalizer, resulting in a single convolute at the point X_0. The box is then moved down one line, and the process repeated

block in the center may be considered to be a separate piece of paper on which are written a new set of abscissa numbers, ranging from -2 thru zero to $+2$. The C's at the right represent the convoluting integers. For the moving average each C is numerically equal to one. To perform a convolution of the ordinate numbers in the table of data with a set of convoluting integers, C_i, each number in the block is multiplied by the corresponding number in the table of data, the resulting products are added and this sum is divided by five. The set of ones is the convoluting function, and the number by which we divide, in this case, 5, is the normalizing factor. To get the next point in the moving average, the center block is slid down one line and the process repeated.

The concept of convolution can be generalized beyond the simple moving average. In the general case the C's represent any set of convoluting integers. There is an associated normalizing or scaling factor. The procedure is to multiply C_{-2} times the number opposite it, then C_{-1} by its number, etc., sum the results, divide by the normalizing factor, if appropriate, and the result is the desired function evaluated at the point indicated by C_0. For the next point, we move the set of convoluting integers down and repeat, etc. The mathematical description of this process is:

$$Y_j^* = \frac{\sum_{i=-m}^{i=m} C_i Y_{j+i}}{N}$$

The index j represents the running index of the ordinate data in the original data table.

For the moving average, each C_i is equal to one and N is the number of convoluting integers. However, for many types of data the set of all 1's, which yields the average, is not particularly useful. For example, on going through a sharp peak, the average would tend to degrade the end of the peak. There are other types of smoothing functions which might be used, and a few of these are indicated in Figure 2.

Figure 2*A* illustrates the set where all values have the same weight over the interval—essentially the moving average.

The function in Figure 2*B* is an exponential set which simulates the familiar *RC* analog time constant—i.e., the most recent point is given the greatest weight, and each preceding point gets a lesser weight determined by the law of exponential decay. *Future* points have no influence. Such a function treats *future* and *past* points differently and so will obviously introduce a unidirectional distortion into the numerical results, as does the *RC* filter in an actual instrument.

When dealing with sets of numbers in hand, and not an actual run on an instrument where the data is emerging in serial order, it is possible to look ahead as well as behind. Then we can convolute with a function that treats *past* and *future* on an equal basis, such as the function in Figure 2*D*. Here the most weight is given to the central point, and points on either side of the center are symmetrically weighed exponentially. This function acts like an idealized lead-lag network, which is not practical to make with resistors, capacitors, and so on.

The usual spectrum from a spectrophotometer is the resultant of two convolutions of the actual spectrum of the material, first with a function representing the slit function of the instrument, which is much like the triangular convolute shown in Figure 2*C*, and then this first convolute spectrum is further convoluted with a function representing the time constant of the instrument. The triangular convoluting function could in many cases yield results not significantly different from the symmetrical exponential function.

Figure 3 illustrates the way in which each of these functions would act on a typical set of spectroscopic data. Curve 3*A* is replotted directly from the instrumental data. It is a single sharp band recorded under conditions which yield a reasonable noise level. The isolated point just to the right of the band has the value of 666 on the scale of zero to 1000 corresponding to approximately 0 to 100% transmittance. This point is introduced to illustrate the effect on these operations of a single point which has a gross error. The numbers along the bottom are the digital value at the

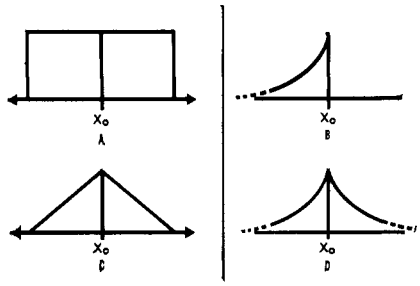

Figure 2. Various convolute functions

A. Moving average. B. Exponential function. C. Symmetrical triangular function, representing idealized spectrometer slit function. D. Symmetrical exponential function

lowest point of the plot, and one may consider that the peak goes down to 34.2% transmittance. The base line at the top is at about 79%.

Curve 3*B* is a nine-point moving average of the data. As expected, the peak is considerably shortened by this process. Especially interesting is the step introduced by the isolated error. In effect, it has the shape of the boxlike convolute in Figure 2*A*, which is exactly what one would expect from the convoluting process (*3*).

Curve 3*C* is for a triangular function which obviously forces both the peak itself and the isolated error into a triangular mold.

Curve 3*D* is the result of convoluting with the numerical equivalent of a conventional *RC* exponential time constant filter using only five points. The peak is not only shortened, but is also shifted to the right by one data point, or 0.002 micron and the isolated data point is asymmetric in the same manner. The convolution with a symmetrical lead-lag exponential, as in Figure 3*E*, does not distort the peak but does still reduce its intensity.

Note that while all of these functions have had the desired effect of reducing the noise level, they are clearly undesirable because of the accompanying degradation of the peak intensity.

METHOD OF LEAST SQUARES

The convoluting functions discussed so far are rather simple and do not extract as much information as is possible. The experimenter, if presented with a plot of the data points, would tend to draw through these points a line which best fits them. Numerically, this can also be done, provided one can adequately define what is meant by best fit. The most common criterion is that of least squares which may be simply stated as follows:

A set of points is to be fitted to some curve—for example, the curve $a_3x^3 + a_2x^2 + a_1x + a_0 = y$. The a's are to be selected such that when each abscissa point is substituted into this equation,

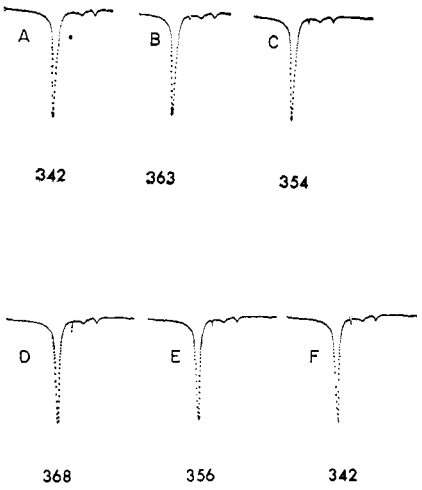

Figure 3. Spectral band convoluted by the various 9 point functions

The number at the bottom of each peak refers to the lowest recorded point, and is a measure of the ability to retain the shape of the peak. *A.* Raw data with single isolated error point. *B.* Moving average. *C.* Triangular function. *D.* Normal exponential function. *E.* Symmetrical exponential function. *F.* Least squares smoothing function

the square of the differences between the computed numbers, y, and the observed numbers is a minimum for the total of the observations used in determining the coefficients. All of the error is assumed to be in the ordinate and none in the abscissa.

Consider the block of seven data points enclosed by the left bracket in Figure 4. If these fall along a curve that can be described approximately by the equation shown, then there are specific procedures—which are described in most books on numerical analysis—to find the a's. One then substitutes back into the resulting equation the abscissa at the central point indicated by the circle. The value which is obtained by this procedure is the best value at that point based on the least squares criterion, on the function which was chosen, and on the group of points examined.

This procedure can be repeated for each group of seven points, dropping one at the left and picking up one at the right each time. A somewhat later block is indicated at the right. In the usual case, there is found a different set of coefficients for each group of seven points. Even with a high-speed computer this is a tedious proposition at best.

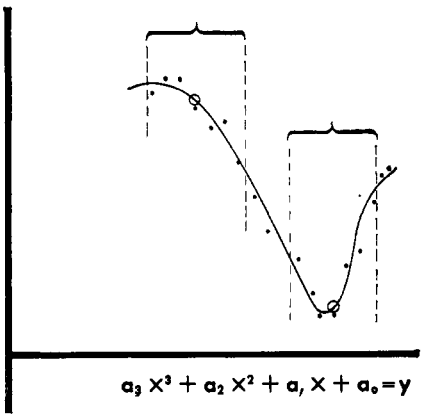

Figure 4. Representation of a 7-point moving polynomial smooth

Note, however, that finding the a coefficients is required only as a means for determining the final best value at just one point, the central point of the set. A careful study of the least squares procedure using these constraints, leads to the derivation of a set of integers which provide a weighting function. With this set of integers the central point can be evaluated by the convoluting procedure discussed above. This procedure is exactly equivalent to the least squares. It is not approximate.

The derivation is presented in Appendix I. For either a cubic or a quadratic function, the set of integers is the same, and the set for up to 25 points is shown in Table I of Appendix II with the appropriate normalizing factors. A most instructive exercise is to tabulate a simple function such as $y = x^3$ over any interval, apply these smoothing convolutes and compare these new values with the original. The answers will be found to be exact.

In Figure 3*F* this least squares convoluting procedure has been applied to the data of Figure 3*A*, using a 9-point cubic convolute. The value at the peak and the shape of the peak are essentially undistorted. As always, the isolated point assumes the shape of the convoluting function. The FORTRAN language computer program for performing this operation is presented in Program I of Appendix III.

Going beyond simple curve fitting, one can find in the literature on numerical analysis a variety of least squares procedures for determining the first derivative. These procedures are usually based on interpolation formulas and are for data at any arbitrary interval. Again, if we restrict ourselves to evaluating the function only at the center point of a set of equally spaced observations, then there exist sets of convoluting integers for the first derivative as well. (These actually evaluate the derivative of the least squares best function.)

A complete set of tables for derivatives up to the fifth order for polynomials up to the fifth degree, using from 5 to 25 points, is presented in Appendix II. These are more than adequate for most work, since, if the points are taken sufficiently close together, then practically any smooth curve will look more or less like a quadratic in the vicinity of a peak, or like a cubic in the vicinity of a shoulder. More complete tables can be found in the statistical literature (*2*, *4*, *6*, *9*).

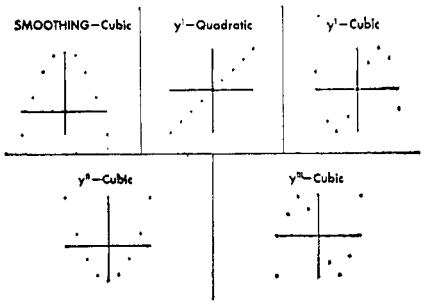

Figure 5. 9-Point convoluting functions (orthogonal polynomial) for smoothing and first, second, and third derivative

Program II of Appendix III shows the use of these tables to obtain the coefficients of a polynomial for finding the precise center of an infrared band.

The shapes of the 9-point convolutes for a few of the functions are illustrated in Figure 5. Of special interest is the linear relation of the first derivative convolute for a quadratic. This is quite unique operationally because in processing a table of data, only one multiplication is necessary for each convolution. The remainder of the points are found from the set calculated for the previous point by simple subtraction. In Figure 6*B* is shown the first derivative of the spectrum in Figure 6*A*, obtained using a 9-point convolute.

The derivatives are useful in cases such as our methods of band finding on a computer (*7*), in studies of derivative spectra, in derivative thermogravimetric analysis, derivative polarography, etc.

CONCLUSIONS

With the increase in the application of computers to the analysis of digitized data, the convolution methods described are certain to gain wider usage. With these methods, the sole function of the computer is to act as a filter to smooth the noise fluctuations and hopefully to introduce no distortions into the recorded data (*3*).

This problem of distortion is difficult to assess. In any of the curves of Figure 3, there remain small fluctuations in the

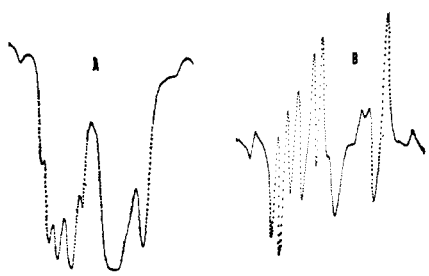

Figure 6. 17-Point first derivative convolute

A. Original spectrum. *B.* First derivative spectrum.

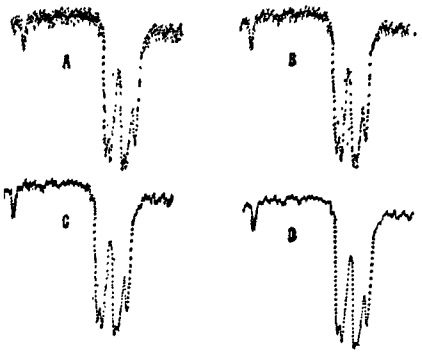

Figure 7. Square root relation between number of points and degree of smoothing

A. Raw data. B. 5-Point smooth. C. 9-Point smooth. D. 17-Point smooth

data. Are these fluctuations real, or, as is more likely, are they just a low frequency component of the noise level which could not be smoothed? The question cannot be answered by taking just the data from a single run. However, if one were to take more than one run, average these and then smooth, or smooth and then average, the computer-plus-instrument system could decide, since even low frequency noise will not recur in exactly the same place in different runs. Computer time is most efficiently used if the averaging is done prior to smoothing.

Recent work (*1*) has shown the utility of simple averaging of a large number of runs in the enhancement of signa-to-noise ratios. The use of combined smoothing and averaging can considerably reduce the instrument time required, throwing the burden onto the computer which operates in a wholly different time domain. A characteristic of both procedures is that the noise is reduced approximately as the square root of the number of points used. This is illustrated for the smoothing case by Figure 7. At the upper left is the raw data, at the upper right a 5-point smooth, lower left 9 points, and lower right 17 points. If it is desired to improve the signal-to-noise ratio by a factor of 10, simple averaging would require a total of 100 runs. Similar improvement could be achieved by making only 16 runs plus a 9-point least squares smooth (average of 16 runs $\leqq 4\times$ improvement, and 9-point smooth $\leqq 3\times$ improvement) or only 4 runs plus a 25-point smooth (average of 4 runs $\leqq 2\times$ improvement plus 25-point smooth $\leqq 5\times$ improvement). The distribution between the number of runs required and the number of points which may be used for the smoothing is a function of the experimental curve under examination. The minimum distortion will occur when the polynomial accurately describes the analytical data, and will deviate as the polynomial departs from the true curve. The best results are obtained when the data are digitized at high densities—i.e., points very close together—and the number of points used in the convolute is chosen to be small enough so that no more than one inflection in the observed data is included in any convolution interval. Our results should be compared with those achievable using conventional instrument filtering. In a sense, we are substituting idealized filters and filter-networks for electronic hardware such as resistors, condensers, servos, etc. If one examines the time relationships, it probably takes longer to get the information using the digitizer and then the computer than with the analogous electronic network. There is, very definitely, the advantage in the computer of being able to vary the processing completely unfettered by the practical restriction of real circuitry and servo loops. Note too, that this processing can be done after the fact of data collection, and indeed several different procedures may be applied in order to assess the optimum. This is a real advantage in itself, and provides ample justification for use of the computer solely as a noise filter. However, the greatest utility of these methods comes in the pretreatment of data to be further processed, as in our bandfinding procedures, in any curvefitting operations, in quantitative analyses, etc.

This type of data processing, so far as computers are concerned, requires a relatively small amount of programming and relatively little use of the computer memory or of the computer's processing capability. Therefore, even accounting-type computers, such as the IBM 1401 can be used to process data in this way. Furthermore, on such computers there is generally a high-speed line printer which can be turned into a relatively crude point plotter. On each line an X can be placed at the appropriate position to 1% of value, and the actual value is printed at the edge of the paper. The rate can be on the order of 10 lines, or points, per second.

ACKNOWLEDGMENT

The authors appreciate the assistance of Harrison J. M. Kinyua in the recompilation and checking of the convolute tables, of Robert Bernard of the Perkin-Elmer Scientific Computing Facility for his advice and assistance in program development, and of Joel Stutman, University of Maryland, for his advice on Appendix I.

APPENDIX I

The general problem is formulated as follows:

A set of $2m + 1$ consecutive values are to be used in the determination of the best mean square fit through these values of a polynomial of degree n (n less than $2m + 1$). This polynomial is of the form(:

$$f_i = \sum_{k=0}^{k=n} b_{nk} i^k = \qquad \text{I}$$

$$b_{n0} + b_{n1} i + b_{n2} i^2 + \ldots + b_{nn} i^n \qquad \text{Ia}$$

The derivatives of this polynomial are:

$$\frac{df_i}{di} = b_{n1} + 2b_{n2} i + 3b_{n3} i^2 + \ldots + nb_{nn} i^{n-1} \qquad \text{Ib}$$

$$\frac{d^2 f_i}{di^2} = 2b_{n2} + 3 \times 2b_{n3} i + \ldots + (n-1) n b_{nn} i^{n-2} \qquad \text{Ic}$$

$$\frac{d^n f_i}{di^n} = n! b_{nn} \qquad \text{Id}$$

Note that, in the coordinate system being considered, the value of i ranges from $-m$ to $+m$, and that $i=0$ at the central point of the set of $2m+1$ values. Hence, the value of the sth derivative at that point is given by:

$$\left(\frac{d^s f_i}{di^s}\right)_{i=0} = s! b_{ns} = a_{ns} \qquad \text{II}$$

where

$$f_0 = b_{n0} = a_{n0} \qquad \text{IIa}$$

$$\frac{df_0}{di} = b_{n1} = a_{n1} \qquad \text{IIb}$$

$$\frac{d^2 f_0}{di^2} = 2b_{n2} = a_{n2} \qquad \text{IIc}$$

The least squares criterion requires that the sum of the squares of the differences between the observed values, y_i, and the calculated, f_i, be a minimum over the interval being considered.

$$\frac{\partial}{\partial b_{nk}} \left[\sum_{i=-m}^{i=m} (f_i - y_i)^2 \right] = 0 \qquad \text{III}$$

Minimizing with respect to b_{n0}, we have

$$\frac{\partial}{\partial b_{n0}} \left[\sum_{i=-m}^{i=m} b_{n0} + b_{n1} i + \ldots + (b_{nn} i^n - y_i)^2 \right] = \qquad \text{IIIa}$$

$$2 \sum_{i=-m}^{i=m} (b_{n0} + b_{n1} i + \ldots + b_{nn} i^n - y_i) = 0$$

and with respect to b_{ni}, we have

$$\frac{\partial}{\partial b_{ni}} \left[\sum_{i=-m}^{i=m} (b_{n0} + b_{n1} i + \ldots + b_{nn} i^n - y_i)^2 \right] = 2 \sum_{i=-m}^{i=m} (b_{n0} + b_{n1} i + \ldots + b_{nn} i^n - y_i)\, i = 0 \qquad \text{IIIb}$$

and with respect to the general b_{nr}, we obtain

$$2\sum_{i=-m}^{i=m}\left[\left(\sum_{k=0}^{k=n} b_{nk}i^k\right) - y_i\right] i^r = 0 \qquad \text{IIIc}$$

or

$$\sum_{i=-m}^{i=m}\sum_{k=0}^{k=n} b_{nk}i^{k+r} = \sum_{i=-m}^{i=m} y_i i^r \qquad \text{IVa}$$

Where r is the index representing the equation number which runs from 0 to n (there are $n + 1$ equations). The summation indexes on the left side may be interchanged—i.e.,

$$\sum_{i=-m}^{i=m}\sum_{k=0}^{k=n} b_{nk}i^{k+r} = \sum_{k=0}^{k=n}\sum_{i=-m}^{i=m} b_{nk}i^{k+r} \qquad \text{IVb}$$

and finally, since b_{nk} is independent of i,

$$\sum_{k=0}^{k=n} b_{nk}\sum_{i=-m}^{i=m} i^{k+r} = \sum_{i=-m}^{i=m} y_i i^r = F_k \qquad \text{Va}$$

or

$$\sum_{k=0}^{k=n} b_{nk}S_{r+k} = F_k \qquad \text{Vb}$$

where

$$S_{r+k} = \sum_{i=-m}^{i=m} i^{r+k} \qquad \text{Vc}$$

and

$$F_k = \sum_{i=-m}^{i=m} i^k y_i \qquad \text{Vd}$$

Note that $S_{r+k} = 0$ for odd values of $r + k$. Since S_{r+k} exists for even values of $r + k$ only, the set of $n + 1$ equations can be separated into two sets, one for even values of k and one for odd values. Thus, for a 5th degree polynomial, where $n = 5$

$$\begin{aligned} S_0b_{50} + S_2b_{52} + S_4b_{54} &= F_0 \\ S_2b_{50} + S_4b_{52} + S_6b_{54} &= F_2 \qquad \text{VIa} \\ S_4b_{50} + S_6b_{52} + S_8b_{54} &= F_4 \end{aligned}$$

which can be used to solve for b_{50}, b_{52}, and b_{54}, while:

$$\begin{aligned} S_2b_{51} + S_4b_{53} + S_6b_{55} &= F_1 \\ S_4b_{51} + S_6b_{53} + S_8b_{55} &= F_3 \qquad \text{VIb} \\ S_6b_{51} + S_8b_{53} + S_{10}b_{55} &= F_5 \end{aligned}$$

which can be used to solve for b_{51}, b_{53}, and b_{55}. The set of equations in VIa has the same form for $n = 4$, so that $b_{40} = b_{50}$, $b_{42} = b_{52}$, and $b_{44} = b_{54}$ while the set VIb has the same form for $n=6$, so that $b_{61} = b_{51}$, $b_{63} = b_{53}$, and $b_{65} = b_{55}$. In other words, $b_{ns} = b_{n+1,s}$ for n and s both even or for n and s both odd. For example, to determine the third derivative for the best fit to a curve of third (or fourth) order, we would have:

$$S_2b_{31} + S_4b_{33} = F_1$$

$$S_4b_{31} + S_6b_{33} = F_3$$

from which $b_{33} = \dfrac{S_2F_3 - S_4F_1}{S_2S_6 - S_4^2}$

When, for instance, $m = 4$ ($2m + 1 = 9$ points), we have from Vc that

$$S_2 = 60,\ S_4 = 708,\ S_6 = 9780$$

and

$$b_{33} = \frac{60F_3 - 708F_1}{60\ (9780) - (708)^2} = \frac{F_3 - 7F_1}{7128}$$

which reduces to:

$$\frac{-14y_{-4} + 7y_{-3} + 13y_{-2} + 9y_{-1} + 0y_0 - 9y_1 - 13y_2 - 3y_3 + 14y_4}{1188}$$

The coefficients of y_i constitute the convoluting integers (Table VIII) for the third derivative of a cubic poly-

TABLE I

CONVOLUTES	SMOOTHING			QUADRATIC	CUBIC		A20	A30			
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	-253										
-11	-138	-42									
-10	-33	-21	-171								
-09	62	-2	-76	-136							
-08	147	15	9	-51	-21						
-07	222	30	84	24	-6	-78					
-06	287	43	149	89	7	-13	-11				
-05	322	54	204	144	18	42	0	-36			
-04	387	63	249	189	27	87	9	9	-21		
-03	422	70	284	224	34	122	16	44	14	-2	
-02	447	75	309	249	39	147	21	69	39	3	-3
-01	462	78	324	264	42	162	24	84	54	6	12
00	467	79	329	269	43	167	25	89	59	7	17
01	462	78	324	264	42	162	24	84	54	6	12
02	447	75	309	249	39	147	21	69	39	3	-3
03	422	70	284	224	34	122	16	44	14	-2	
04	387	63	249	189	27	87	9	9	-21		
05	322	54	204	144	18	42	0	-36			
06	287	43	149	89	7	-13	-11				
07	222	30	84	24	-6	-78					
08	147	15	9	-51	-21						
09	62	-2	-76	-136							
10	-33	-21	-171								
11	-138	-42									
12	-253										
NORM	5175	8059	3059	2261	323	1105	143	429	231	21	35

TABLE II

CONVOLUTES	SMOOTHING			QUARTIC		QUINTIC		A40	A50		
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	1265										
-11	-345	285									
-10	-1122	-114	11628								
-09	-1255	-285	-6460	340							
-08	-915	-285	-13005	-255	1195						
-07	-255	-165	-11220	-420	-1 95	2145					
-06	590	30	-3940	-290	-2 60	-2860	110				
-05	1503	261	6378	18	-1 17	-2937	-198	18			
-04	2385	495	17655	405	1 35	-165	-160	-45	15		
-03	3155	705	28190	790	4 15	3755	110	-10	-55	5	
-02	3750	870	36660	1110	66 0	7500	390	60	30	-30	
-01	4125	975	42120	1320	82 5	10125	600	120	135	75	
00	4253	1011	44003	1393	88 3	11053	677	143	179	131	
01	4125	975	42120	1320	82 5	10125	600	120	135	75	
02	3750	870	36660	1110	66 0	7500	390	60	30	-30	
03	3155	705	28190	790	41 5	3755	110	-10	-55	5	
04	2385	495	17655	405	135	-165	-160	-45	15		
05	1503	261	6378	18	-117	-2937	-198	18			
06	590	30	-3940	-290	-260	-2860	110				
07	-255	-165	-11220	-420	-195	2145					
08	-915	-285	-13005	-255	195						
09	-1255	-285	-6460	340							
10	-1122	-114	11628								
11	-345	285									
12	1265										
NORM	30015	6555	260015	7429	4199	46189	2431	429	429	231	

TABLE III

CONVOLUTES	1ST DERIVATIVE			QUADRATIC			A21				
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	-12										
-11	-11	-11									
-10	-10	-10	-10								
-09	-9	-9	-9	-9							
-08	-8	-8	-8	-8	-8						
-07	-7	-7	-7	-7	-7	-7					
-06	-6	-6	-6	-6	-6	-6	-6				
-05	-5	-5	-5	-5	-5	-5	-5	-5			
-04	-4	-4	-4	-4	-4	-4	-4	-4	-4		
-03	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	
-02	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
-01	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
00	0	0	0	0	0	0	0	0	0	0	0
01	1	1	1	1	1	1	1	1	1	1	1
02	2	2	2	2	2	2	2	2	2	2	2
03	3	3	3	3	3	3	3	3	3	3	
04	4	4	4	4	4	4	4	4	4		
05	5	5	5	5	5	5	5	5			
06	6	6	6	6	6	6	6				
07	7	7	7	7	7	7					
08	8	8	8	8	8						
09	9	9	9	9							
10	10	10	10								
11	11	11									
12	12										
NORM	1300	1012	770	570	408	280	182	110	60	28	10

TABLE IV

CONVOLUTES	1ST DERIVATIVE			CUBIC		QUARTIC		A31	A41			
POINTS	25	23	21	19	17	15	13	11	9	7	5	
-12	30866											
-11	8602	3938										
-10	-8525	815	84075									
-09	-20982	-1518	10032	6936								
-08	-29236	-3140	-43284	68	748							
-07	-33754	-4130	-78176	-4648	-98	12922						
-06	-35003	-4567	-96947	-7481	-643	-4121	1133					
-05	-33450	-4530	-101900	-8700	-930	-14150	-660	300				
-04	-29562	-4098	-95338	-8574	-1002	-18334	-1578	-294	86			
-03	-23806	-3350	-79564	-8179	-902	-17842	-1796	-532	-142	22		
-02	-16649	-2365	-56881	-5363	-673	-13843	-1489	-503	-193	-67	1	
-01	-8558	-1222	-29592	-2816	-358	-7506	-832	-296	-126	-58	-8	
00	0	0	0	0	0	0	0	0	0	0	0	
01	8558	1222	29592	2816	358	7506	832	296	129	58	8	
02	16649	2365	56881	5363	673	13843	1489	503	193	67	-1	
03	23806	3350	79504	8179	902	17842	1796	532	142	-22		
04	29562	4098	95338	8574	1002	18334	1578	294	-86			
05	33450	4530	101900	8700	930	14150	660	-300				
06	35003	4567	96947	7481	643	4121	-1133					
07	33754	4130	78176	4648	98	-12922						
08	29236	3140	43284	-68	-748							
09	20982	1518	-10032	-6936								
10	8525	-815	84075									
11	-8602	-3938										
12	-30866											
NORM	1776060	197340	3634092	255816	23256	334152	24024	5148	1188	252	12	

TABLE V

CONVOLUTES	1ST DERIVATIVE			QUINTIC		SEXIC		A51	A61			
POINTS	25	23	21	19	17	15	13	11	9	7	5	
-12	-6356625											
-11	-11820675	-357045										
-10	-15593141	-654687	-15977364									
-09	-17062146	-840937	-28754154	-332684								
-08	-15896511	-878634	-35613829	-583549	-23945							
-07	-12139321	-752859	-34807914	-686099	-40483	-175125						
-06	-6301491	-478349	-26040033	-604484	-43973	-279975	-31380					
-05	544668	-106911	-10949942	-348823	-32306	-266401	-45741	-3084				
-04	6671883	265164	6402438	9473	-8671	-130506	-33511	-3776	-5758			
-03	9604353	489687	19052988	322378	16679	65229	-12	-1244	-4538	-90		
-02	6024183	359157	16649358	349928	24661	169819	27093	2166	2762	18		
-01	-8322182	-400653	-15033066	-255102	-14404	-78351	-14647	-573	-508	-2		
00	0	0	0	0	0	0	0	0	0	0		
01	8322182	400653	15033066	255102	14404	78351	14647	573	508	2		
02	-6024183	-359157	-16649358	-349928	-24661	-169819	-27093	-2166	-2762	-18		
03	-9604353	-489687	-19052988	-322378	-16679	-65229	12	1244	4538	90		
04	-6671883	-265164	-6402438	-9473	8671	130506	33511	3776	5758			
05	-544668	106911	10949942	348823	32306	266401	45741	3084				
06	6301491	478349	26040033	604484	43973	279975	31380					
07	12139321	752859	34807914	686099	40483	175125						
08	15896511	878634	35613829	583549	23945							
09	17062146	840937	28754154	332684								
10	15593141	654687	15977364									
11	11820675	357045										
12	6356625											
NORM	7153575	312455	5311735	81719	41990	20995	2431	143	143	1		

TABLE VI

CONVOLUTES	2ND DERIVATIVE			QUADRATIC	CUBIC		A22	A32			
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	92										
-11	69	77									
-10	48	56	190								
-09	29	37	133	51							
-08	12	20	82	34	40						
-07	-3	5	37	19	25	91					
-06	-16	-8	-2	6	12	52	22				
-05	-27	-19	-35	-5	1	19	11	15			
-04	-36	-28	-62	-14	-8	-8	2	6	28		
-03	-43	-35	-83	-21	-15	-29	-5	-1	7	5	
-02	-48	-40	-98	-26	-20	-48	-10	-6	-8	0	2
-01	-51	-43	-107	-29	-23	-53	-13	-9	-17	-3	-1
00	-52	-44	-110	-30	-24	-56	-14	-10	-20	-4	-2
01	-51	-43	-107	-29	-23	-53	-13	-9	-17	-3	-1
02	-48	-40	-98	-26	-20	-48	-10	-6	-8	0	-2
03	-43	-35	-83	-21	-15	-29	-5	-1	7	5	
04	-36	-28	-62	-14	-8	-8	2	6	28		
05	-27	-19	-35	-5	1	19	11	15			
06	-16	-8	-2	6	12	52	22				
07	-3	5	37	19	25	91					
08	12	20	82	34	40						
09	29	37	133	51							
10	48	56	190								
11	69	77									
12	92										
NORM	26910	17710	33649	6783	3876	6188	1001	429	462	42	7

TABLE VII

CONVOLUTES	2ND DERIVATIVE			QUARTIC		QUINTIC		A42	A52		
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	-429594										
-11	31119	-346731									
-10	298155	61845	-37791								
-09	413409	281979	11628	-96084							
-08	414786	358530	35802	45084	-121524						
-07	336201	331635	41412	105444	82251	-93093					
-06	207579	236709	34353	109071	153387	88803	-72963				
-05	54855	104445	19734	76830	137085	133485	98010	-10530			
-04	-100026	-39186	1878	26376	71592	95568	115632	20358	-4158		
-03	-239109	-172935	-15678	-27846	-11799	19737	53262	17082	12243	-117	
-02	-348429	-280275	-30183	-74601	-88749	-59253	-32043	117	4983	603	-3
-01	-418011	-349401	-39672	-105864	-141873	-116577	-99528	-15912	-6963	-171	48
00	-441870	-373230	-42966	-116820	-160740	-137340	-124740	-22230	-12210	-630	-90
01	-418011	-349401	-39672	-105864	-141873	-116577	-99528	-15912	-6963	-171	48
02	-348429	-280275	-30183	-74601	-88749	-59253	-32043	117	4983	603	-3
03	-239109	-172935	-15678	-27846	-11799	19737	53262	17082	12243	-117	
04	-100026	-39186	1878	26376	71592	95568	115632	20358	-4158		
05	54855	104445	19734	76830	137085	133485	98010	-10530			
06	207579	236709	34353	109071	153387	88803	-72963				
07	336201	331635	41412	105444	82251	-93093					
08	414786	358530	38802	45084	121524						
09	413409	281979	11628	-96084							
10	298155	61845	-37791								
11	31119	-346731									
12	429594										
NORM	4292145	2812095	245157	490314	478686	277134	160446	16731	4719	99	3

TABLE VIII

CONVOLUTES	3RD DERIVATIVE			CUBIC	QUARTIC		A33	A43			
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	-506										
-11	-253	-77									
-10	-55	-35	-285								
-09	93	-3	-114	-204							
-08	196	20	12	-68	-28						
-07	259	35	98	28	-7	-91					
-06	287	43	149	89	7	-13	-11				
-05	285	45	170	120	15	35	0	-30			
-04	258	42	166	126	18	58	6	6	-14		
-03	211	35	142	112	17	61	8	22	7	-1	
-02	145	25	103	83	13	49	7	23	13	1	-1
-01	77	13	54	44	7	27	4	14	9	1	2
00	0	0	0	0	0	0	0	0	0		0
01	-77	-13	-54	-44	7	-27	-4	-14	-9	-1	-2
02	-149	-25	-103	-83	13	-49	-7	-23	-13	-1	1
03	-211	-35	-142	-112	17	-61	-8	-22	-7	1	
04	-258	-42	-166	-126	18	-58	-6	-6	14		
05	-285	-45	-170	-120	15	-35	0	30			
06	-287	-43	-149	-89	7	13	11				
07	-259	-35	-98	-28	-7	91					
08	-196	-20	-12	68	28						
09	-93	3	114	204							
10	55	35	285								
11	253	77									
12	506										
NORM	296010	32890	86526	42636	3876	7956	572	858	198	6	2

TABLE IX

CONVOLUTES	3RD DERIVATIVE			QUINTIC	SEXIC		A53	A63			
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	118745										
-11	217640	23699									
-10	279101	42704	425412								
-09	290076	52959	749372	317655							
-08	244311	51684	887137	1113240	4915						
-07	144616	38013	787382	1231500	8020	93135					
-06	5131	13632	448909	932760	7975	141320	11260				
-05	-146408	-16583	-62644	259740	4380	113065	15250	1580			
-04	-266403	-43928	-598094	-589080	-1755	3800	8165	1700	2295		
-03	-293128	-55233	-908004	-1220520	-7540	-150665	-6870	-55	1280	65	
-02	-144463	-32224	-625974	-1007760	-7735	-260680	-16335	-2010	-2285	-40	
-01	284372	49115	748068	948600	5720	-169295	7150	645	500	5	
00	0	0	0	0	0	0	0	0	0	0	
01	-284372	-49115	-748068	-948600	-5720	-169295	-7150	-645	-500	-5	
02	144463	32224	625974	1007760	7735	260680	16335	2010	2285	40	
03	293128	55233	908004	1220520	7540	150665	6870	55	-1280	-65	
04	266403	43928	598094	589080	1755	-3800	-8165	-1700	-2295		
05	146408	16583	62644	-259740	-4380	-113065	-15250	-1580			
06	-5131	-13632	-448909	-932760	-7975	-141320	-11260				
07	-144616	-38013	-787382	-1231500	-8020	-93135					
08	-244311	-51684	-887137	-1113240	-4915						
09	-290076	-52959	-749372	-317655							
10	-279101	-42704	-425412								
11	-217640	-23699									
12	-118745										
NORM	5722860	749892	4249388	4247012	16796	2144809	9724	572	286	2	

TABLE X

CONVOLUTES	4TH DERIVATIVE			QUARTIC		QUINTIC		A44	A54		
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	858										
-11	803	858									
-10	643	793	594								
-09	393	605	540	396							
-08	78	315	385	352	36						
-07	-267	-42	150	227	31						
-06	-597	-417	-130	42	17	621	84				
-05	-857	-747	-406	-168	-3	251	64	6			
-04	-982	-955	-615	-354	-24	-249	11	4	18		
-03	-897	-950	-680	-453	-39	-704	-54	-1	9	6	
-02	-517	-627	-510	-388	-39	-869	-96	-6	-11	1	
-01	253	133	0	-68	-13	-429	-66	-6	-21	-7	
00	1518	1463	969	612	52	1001	99	6	14	-3	
01	253	133	0	-68	-13	-429	-66	-6	-21	-7	
02	-517	-627	-510	-388	-39	-869	-96	-6	-11	1	
03	-897	-950	-680	-453	-39	-704	-54	-1	9	6	
04	-982	-955	-615	-354	-24	-249	11	4	18		
05	-857	-747	-406	-168	-3	251	64	6			
06	-597	-417	-130	42	17	621	84				
07	-267	-42	150	227	31	756					
08	78	315	385	352	36						
09	393	605	540	396							
10	643	793	594								
11	803	858									
12	858										
NORM	1430715	937365	408595	163438	8398	92378	4862	143	143	11	

TABLE XI

CONVOLUTES	5TH DERIVATIVE			QUINTIC		SEXIC		A55	A65		
POINTS	25	23	21	19	17	15	13	11	9	7	5
-12	-275										
-11	-500	-65									
-10	-631	-116	-1404								
-09	-636	-141	-2444	-44							
-08	-501	-132	-2819	-74	-55						
-07	-236	-87	-2354	-79	-88	-675					
-06	119	-12	-1063	-54	-83	-1000	-20				
-05	488	77	788	-3	-36	-751	-26	-4			
-04	753	152	2618	58	39	44	-11	-4	-9		
-03	748	171	3468	98	104	979	18	1	-4	-5	
-02	253	76	1938	68	91	1144	33	6	11	4	
-01	-1012	-209	-3876	-102	-104	-1001	-22	-3	-4	-1	
00	0	0	0	0	0	0	0	0	0	0	
01	1012	209	3876	102	104	1001	22	3	4	1	
02	-253	-76	-1938	-68	-91	-1144	-33	-6	-11	-4	
03	-748	-171	-3468	-98	-104	-979	-18	-1	4	5	
04	-753	-152	-2618	-58	-39	-44	11	4	9		
05	-488	-77	-788	3	36	751	26	4			
06	-119	12	1063	54	83	1000	20				
07	236	87	2354	79	88	675					
08	501	132	2819	74	55						
09	636	141	2444	44							
10	631	116	1404								
11	500	65									
12	275										
NORM	1300650	170430	1931540	29716	16796	83980	884	52	26	2	

PROGRAM 1

```
*       SUBROUTINE SMOOTH -9 POINT                                          1
C                                                                           2
        SUBROUTINE SMOOTH (N,NDATA,M,MDATA)                                 3
C                                                                           4
C       INPUTS                                                              5
C          N       NUMBER OF RAW DATA POINTS                                6
C          NDATA   ARRAY OF N RAW DATA POINTS STORED IN MAIN PROGRAM         7
C                     DUMMY DIMENSION                                       8
C       OUTPUTS                                                             9
C          M       NUMBER OF SMOOTHED DATA POINTS = N-8                     10
C          MDATA   ARRAY OF M SMOOTHED POINTS STORED IN MAIN PROGRAM        11
C                     MAY BE SAME REGION IN MAIN PROGRAM AS NDATA           12
C                     DUMMY DIMENSION                                       13
C                                                                           14
        DIMENSION NDATA(1000),MDATA(1000),NP(9)                             15
C                                                                           16
C       INITIALIZATION SEGMENT                                              17
C                                                                           18
        M=N-8                                                               19
        DO 10 I=2,9                                                         20
        J=I-1                                                               21
   10   NP(I) = NDATA(J)                                                    22
C                                                                           23
C       SMOOTHING LOOP                                                      24
C                                                                           25
        DO 200 I=1,M                                                        26
        J=I+8                                                               27
        DO 11 K=1,8                                                         28
        KA = K+1                                                            29
   11   NP(K)=NP(KA)                                                        30
        NP(9) = NDATA(J)                                                    31
        NSUM=59*NP(5)+54*(NP(4)+NP(6))+39*(NP(3)+NP(7))+14*(NP(2)+NP(8))-   32
       121*(NP(1)+NP(9))                                                    33
        MDATA(I) = NSUM/231                                                 34
   200  CONTINUE                                                            35
C                                                                           36
  9999  RETURN                                                              37
C                                                                           38
*       END                                                                 39
```

nomial determined from a least squares fit to 9 points. Since the value of a_{33} is $3!b_{33}$, the denominator in the above expression must be divided by 6 to get the normalizer of 198 found in Table III.

In all of the above derivations, it has been assumed that the sampling interval is the same as the absolute abscissa interval—i.e., $\Delta x = 1$. If not, the value of Δx must be included in the normalization procedure. Hence, to evaluate the sth derivative at the central point of a set of m values, based on an nth degree polynomial fit, we must evaluate

$$a_{nsm} = s!b_{nsm} = \frac{\sum_{i=-m}^{i=m} C_{ism} y_i}{\Delta x^s N_{sm}} \qquad \text{VII}$$

Note that since $\Delta x^0 = 1$, the interval is of no concern in the case of smoothing.

Repeated Convolution. The process of convolution can be repeated if desired. For example, one might wish to further smooth a set of previously smoothed points, or to obtain the derivative only after the raw data has been smoothed. Thus, if we convolute using p points the first time, and m points the second,

$$(f)_{s_1 s_2} = a_{n_2 s_2 m} = \frac{\sum_{i=-m}^{i=m} C_{is_2 m}\, a_{n_1 s_1 p}}{\Delta x^{(s_2)} N_{s_2 m}} \qquad \text{VIII}$$

$$= \frac{\sum_{i=-m}^{i=m} C_{is_2 m} \sum_{j=-p}^{j=p} C_{js_1 p} F_j}{\Delta x^{(s_1+s_2)} N_{s_2 m} N_{s_1 p}} \qquad \text{IX}$$

$$= \frac{\sum_{i=-m}^{i=m} \sum_{j=-p}^{j=p} C_{is_2 m}\, C_{js_1 p}\, F_{i+j}}{\Delta x^{(s_1+s_2)} N_{s_2 m} N_{s_1 p}} \qquad \text{X}$$

thus:

$$(f)_{s_1 s_2} = \frac{\sum_{h=-(m+p)}^{h=(m+p)} d_h y_h}{N_h} \qquad \text{XI}$$

where $h = i + j$

$$N_h = \Delta x^{(s_1 s_2)} N_{s_1 p} N_{s_2 m}$$

$$d_h = \sum_{i=-m}^{i=m} \sum_{j=-p}^{j=p} c_i c_j$$

Equation XI shows that one need not go through the convolution procedure twice, but can do a single convolution, using $2(m+p)+1$ points, and a table of new integers formed by combining the c's.

For the case where a cubic smooth is to be followed by obtaining the quadratic first derivative using $m=2$ and $p=2$:

$C_i = -2, -1, 0, 1, 2, N_i = 10$
$C'_j = -3, 12, 17, 12, -3, N_j = 35$
$d_{-4} = C_{-2}C'_{-2} = 6$
$d_{-3} = C_{-2}C'_{-1} + C_{-1}C'_{-2} = -36 + 3 = -33$
$d_{-2} = C_{-2}C'_0 + C_0C'_2 + C_{-1}C'_{-1} = -34 - 12 = -46$
$d_{-1} = C_{-2}C'_1 + C_1C'_{-2} + C_{-1}C'_0 + C_0C'_{-1} = -24 - 3 - 17 = -44$
$d_0 = C_{-2}C'_2 + C_2C'_{-2} + C_1C'_{-1} - C_{-1}C'_1 + C_0C'_0 = 0$
$d_1 = \text{By symmetry} = 44$
$d_2 = 46$
$d_3 = 33$
$d_4 = -6$
$N_h = 350\Delta x$

PROGRAM 2

```
*     SUBROUTINE CENTER    LORENZ
C                                                                        2
      SUBROUTINE CENTER(NPOINT,X,Y,TRANS,DENS)                           3
C                                                                        4
C     COMPUTATION OF PRECISE PEAK POSITION AND INTENSITY USING 9 POINTS  5
C     TO A QUADRATIC,Y=A20+A21X+A22XSQ, X=(-A21/2A22). IN ORDER TO       6
C     APPROXIMATE A LORENTZ CONTOUR, VALUES ARE CONVERTED TO ABSORBANCE  7
C     AND THE RECIPROCALS ARE USED IN DETERMINING THE COEFFICIENTS BY    8
C     ORTHOGONAL POLYNOMIALS                                             9
C                                                                       10
      DIMENSION NPOINT(25),DNS(9)                                       11
   60 DO 61 I=1,9                                                       12
      IP=I+8                                                            13
   61 DNS(I)=1./ALOG10F(1000./NPOINT(IP))                               14
      P4=DNS(1)+DNS(9)                                                  15
      P3=DNS(2)+DNS(8)                                                  16
      P2=DNS(3)+DNS(7)                                                  17
      P1=DNS(4)+DNS(6)                                                  18
      A20=(-21.*P4)+(14.*P3)+(39.*P2)+(54.*P1)+(59.*DNS(5))             19
      A20=A20/231.                                                      20
      A21=(4.*(DNS(9)-DNS(1)))+(3.*(DNS(8)-DNS(2)))+(2.*(DNS(7)-DNS(3)))21
     1+(DNS(6)-DNS(4))                                                  22
      A21=A21/60.                                                       23
      A22=(28.*P4)+(7.*P3)-(8.*P2)-(17.*P1)-(20.*DNS(5))                24
      A22=A22/924.                                                      25
      X=(-A21/(2.0*A22))                                                26
      Y=A20+X*(A21+X*A22)                                               27
      DENS=1.0/Y                                                        28
      TRANS=1000./(10.0**DENS)                                          29
 1000 RETURN                                                            30
*     END                                                               31
```


APPENDIX II

The following eleven tables contain the convoluting integers for smoothing (zeroth derivative) through the fifth derivative for polynomials of degree two through five. They are in the form of tables of A_{ij}, where i is the degree of the polynomial and j is the order of the derivative. Thus, to obtain the third derivative over 17 points, assuming a fourth degree polynomial (A_{43}), one would use the integers in the column headed 17 of Table VIII.

APPENDIX III. COMPUTER PROGRAMS

The programming of today's high speed digital computers is still an art rather than a science. Different programmers presented with the same problem will, in general, write quite different programs to satisfactorily accomplish the calculation. Two programs are presented here as examples of the techniques discussed in this paper. They are written in the FORTRAN language, since this is one of the most widespread of the computer programming languages. Programmers using other languages should be able to follow the logic quite readily and make the appropriate translations. Each is written as a subroutine for incorporation into a larger program as required.

Program 1 is a 9-point least squares smooth of spectroscopic data. The raw data has previously been stored by the main program in the region NDATA.

Lines 1 through 25 are explanatory and housekeeping to set up the initial conditions. The 9-point array NP contains the current set of data points to be smoothed. The main loop consists of lines 26 through 35. The inner loop, lines 28 through 30, moves the previous set of points up one position. The next point is added by line 31. In lines 32 and 33 the convoluting integers are multiplied by the corresponding data, and the products summed. In line 34, the sum is divided by the normalizing constant and the resulting smoothed point is stored.

Program 2 computes the precise peak position and intensity of a set of points which is known to contain a spectroscopic peak. In order to approximate a Lorentz contour (*5*), values are converted to absorbance and the reciprocals are used in determining the coefficients of a polynomial (*6*) having the form:

$$y = a_{20} + a_{21}x + a_{22}x^2$$

The center is at the point where $x = -a_{21}/2a_{22}$.

The data points to be used are points 9 through 17 stored in the array NPOINT and may have any value from 30 through 999. In the loop lines 12 through 14, each of these points is converted to absorbance, the reciprocal taken, and the result stored in the array DNS.

Since for a_{20} and a_{22}, the convolute function is symmetric about the origin, forming the sums $P4$ through $P1$ in lines 15 through 18 shortens the computation. The first constant, a_{20} is found in line 19, using the values of the 9-point convolute from Table I, and normalized in line 20.

The value of a_{21} is computed in lines 21–22 using the convolute from Table III (first derivative-quadratic). Note that this is an antisymmetric function. Table VI furnished the constants for the computation of a_{22} in line 24. Note that the normalizing factor of 924 is 2! times the value given in the table. X is computed in line 26 and the corresponding value of y is computed in line 27 by substituting the appropriate values into the polynomial. The absorbance or optical density (DENS) is, of course, the reciprocal of y (line 28).

LITERATURE CITED

(1) Allen, L. C., Johnson, L. F., *J. Am. Chem. Soc.* **85,** 2668 (1963).
(2) Anderson, R. L., Houseman, E. F., "Tables of Orthogonal Polynominal Values Extended to N = 104," Agricultural Experiment Station, Iowa State College of Agriculture and Mechanical Arts, Ames, Iowa, 1942.
(3) Cole, LaMont C., *Science* **125,** 874 (1957).
(4) Guest, P. G., "Numerical Methods of Curve Fitting," p. 349 ff., University Press, Cambridge, England, 1961.
(5) Jones, R. N., Seshadri, K. S., Hopkins, J. W., *Can. Jour. Chem.* **40,** 334 (1962).
(6) Kerawala, S. M., *Indian J. Phys.* **15,** 241 (1941).
(7) Savitzky, A., ANAL. CHEM. **33,** 25A (Dec. 1961).
(8) Whitaker, S., Pigford, R. L., *Ind. Eng. Chem.* **52,** 185 (1960).
(9) Whittaker, E., Robinson, G., "The Calculus of Observations," p. 291 ff., Blackie and Son, Ltd., London & Glasgow, 1948.

RECEIVED for review February 10, 1964. Accepted April 17, 1964. Presented in part at the 15th Annual Summer Symposium on Analytical Chemistry, University of Maryland, College Park, Md., June 14, 1962.

Controlled-Potential Coulometric Analysis of N-Substituted Phenothiazine Derivatives

F. HENRY MERKLE and CLARENCE A. DISCHER

College of Pharmacy, Rutgers—The State University, Newark 4, N. J.

▶ **Controlled-potential electrolysis is suitable for the coulometric determination of several pharmaceutically important N-substituted phenothiazines. The concentration of sulfuric acid, used as the supporting electrolyte, has a differentiating effect on the half-wave potentials of the compounds studied. Polarographic measurements obtained with a rotating platinum microelectrode established current-voltage curves. The compounds could be quantitatively oxidized to a free radical or to a sulfoxide by selection of suitable acid concentrations and applied potentials. Electroreduction of the free radicals occurs at approximately +0.25 volt *vs.* S.C.E. on a platinum electrode. In the case of the sulfoxides, a single 2-electron reduction step occurred at ca. −0.95 volt *vs.* S.C.E. on a mercury pool cathode. The determinations showed good reproducibility and an accuracy of ca. 1% was obtained with sample concentrations of 10^{-3} M or greater.**

THE establishment of an oxidation mechanism for the various *N*-amino–substituted phenothiazines is a subject of considerable biological significance (*2*). The character of this oxidation has been investigated by using a variety of analytical methods (*1, 4, 7, 8*). However, because of the transient nature of a free radical intermediate, relatively little has been done to demonstrate the quantitative aspects of this reaction.

The application of controlled-potential electrolysis to the analysis of phenothiazine derivatives is an extension of previous work by the authors on the electro-oxidation of chlorpromazine (*6*). Since these compounds are electrolytically active at moderate applied potentials, controlled-potential coulometry offers a rapid and absolute approach for their quantitative determination. Moreover, this technique makes possible the direct determination of the individual species involved in the oxidation sequence.

The oxidation reactions for phenothiazine derivatives are, in general, represented by:

$$R: \rightarrow R\cdot + e^- \qquad (1)$$

and

$$R\cdot + H_2O \rightarrow S + 2H^+ + e^- \qquad (2)$$

in 12*N* sulfuric acid, and

$$R: \rightarrow R\cdot + e^- \qquad (3)$$

and

$$2R\cdot + H_2O \xrightarrow{\text{spontaneous}} R: + S + 2H^+ \qquad (4)$$

in 1*N* sulfuric acid (*2, 6*), where R: represents the initial reduced form of the compound, R· represents the free radical obtained upon 1-electron oxidation, and S represents the corresponding sulfoxide.

EXPERIMENTAL

Instrumentation. Polarograms were obtained on a Sargent Model XXI recording polarograph. An H-type cell was used, with a sintered-glass disk of medium porosity separating the two electrode compartments. A rotating platinum microelectrode served as the working electrode *vs.* S.C.E. as reference. To obtain well defined reproducible S-shaped curves it was necessary to pretreat the microelectrode by anodic polarization for 10 minutes in 1*N* or 12*N* sulfuric acid at +1.0 volt *vs.* S.C.E., followed immediately by a brief 2- to 3-minute electrolysis with the platinum microelectrode as the cathode.

The controlled-potential electrolyses were performed with an electronic controlled-potential coulometric titrator, Model Q-2005 ORNL (*5*). Readout voltages were measured with a Non-Linear Systems Model 484 A digital voltmeter.

Electrolysis Cells and Electrodes. Two closed cells with operating capacities of 100 and 20 ml., respectively, were used for oxidations. They were constructed to accommodate large cylindrical, wire mesh, rotating platinum electrodes, 2.5 cm. in diameter and 5 cm. in height for the large cell, and 1 cm. in diameter and 3 cm. in height for the small cell. The reference electrode (S.C.E.) and auxiliary platinum cathode were isolated from the sample compartment by sintered-glass diaphragms and agar plugs.

The sample compartment of the reduction cell consisted of a 125-ml. wide-mouthed Erlenmeyer flask with a standard-taper neck and fitted cover. Approximately 30 ml. of mercury was used as the cathode pool. A glass propeller-type stirrer served to agitate the mercury pool. The reference electrode (S.C.E.) and auxiliary electrode (platinum anode) chambers were separated from the cathode compartment by glass side-arms fitted with sintered-glass diaphragms and agar plugs.

The three cells were constructed so that nitrogen gas could be bubbled through the solution before and during electrolysis. The platinum gauze electrodes and the glass stirrer were rotated at 600 r.p.m. using a Sargent Synchronous Rotator.

Materials. REAGENTS. The acid solutions were prepared using sulfuric acid, Baker analyzed reagent, distilled water, and ethanol U.S.P. grade.

PHENOTHIAZINE DERIVATIVES. The compounds studied were provided in powdered form by the suppliers indicated: chlorpromazine hydrochloride, chlorpromazine sulfoxide hydrochloride, prochlorperazine ethanedisulfonate, and trifluoperazine dihydrochloride, Smith, Kline & French Laboratories; promethazine hydrochloride and promazine hydrochloride, Wyeth Laboratories; triflupromazine hydrochloride, E. R. Squibb Laboratories; thioridazine hydrochloride, Sandoz Laboratories. The structural formulas and generic names of the compounds used in this work are shown in Figure 1.