by attenuation. The general form of the equation would then become

$$\theta = 1 - \frac{X_v H_v(d_{Rl} - d_{Rs})}{|d_{Rv} - d_{Rs}|(X_l H_l - X_s H_s) + X_s H_s(d_{Rl} - d_{Rs})} \quad (8)$$

The linearity of the detection system including the attenuator will directly affect the accuracy of θ and should be considered when making calculations.

Using Equation 8, resolution may be calculated for adjacent peaks that may vary in height as much as the dynamic range of the instrument. All that is required to make an accurate determination of θ is the peak and valley heights and their respective retention distances. More information is required to calculate θ in this derived form, but each measurement is relatively easy to make and awkward construction lines are entirely eliminated. This method could easily be applied to systems using computers for data reduction. Those computer systems using continuous sampling techniques would, in fact, acquire all necessary data and need only sufficient programming to make the calculation.

DEFINITION OF SYMBOLS

R	= Resolution (*1*).
θ	= Resolution (*2*).
Subscript i,j	= solute designation: solute i, relative to solute j eluted later.
Subscript l,s	= solute designation: large peak, small peak.
Subscript v	= designates valley between adjacent peaks.
$H_{l,s}$	= height of respective peak.
H_v	= distance from base line to valley between adjacent peaks.
$d_{Ri,j}$, $d_{Rl,s}$	= retention distance for respective peaks.
$Y_{dj,i}$	= width of respective peaks.
d_{Rv}	= retention distance to valley.
$X_{l,s,v}$	= attenuation factor for respective peaks and valley.

GLENN C. CARLE

Ames Research Center, NASA
Moffett Field, Calif. 94035

RECEIVED for review February 2, 1972. Accepted May 12, 1972.

Comments on Smoothing and Differentiation of Data by Simplified Least Square Procedure

SIR: Some years ago, Savitzky and Golay (*1*) proposed a series of numerical tables for the smoothing of experimental data and for the computation of their derivatives. The main advantage of the method is that they obtain universal numerical functions, the convolution of which with the original vector data gives a smoothed vector as well as its successive derivatives.

Unfortunately the published tables seem to include some numerical errors. Moreover, Equation IIIc in (*1*) fails in the case $i = k = r = 0$ since it does not give the identification $b_{n,0} = y_0$ as it follows from Equation IIIa. The correct equation is

$$\sum_{\substack{i=-m \\ i\neq 0}}^{i=+m} \left[\left(\sum_{k=0}^{k=n} b_{n,k} \cdot i^k \right) - y_i \right] i^r + b_{n,0} - y_0 = 0$$

The tables were recalculated by introducing this correction in Savitzky and Golay's reasoning and a check was performed by using a more general matrician formalism (*2, 3*).

Let **Y** be the $(n \times 1)$ vector of observations and ϵ the $(n \times 1)$ vector of "error" random variables with

$$E(\epsilon) = 0 \quad (1)$$

and dispersion matrix

$$V(\epsilon) = E(\epsilon\epsilon') = \sigma^2\mathbf{I} \quad (2)$$

where **I** is the $(n \times n)$ identity matrix.

The relations 1 and 2 imply that the ϵ_i are uncorrelated, they all have zero mean value and the same variance σ^2.

(1) A. Savitzky and M. J. E. Golay, ANAL. CHEM., **36**, 1627 (1964)
(2) J. Steinier, *Bull. Rech. Agron. Gembloux*, in press.
(3) M. Kendall and A. Stuart, "The Advanced Theory of Statistics," Vol. 2, Charles Griffin et G. Ltd., London, 1958–1966, pp 75–87.

Table I. Corrections of Values given by Savitzky and Golay

Table[a]	Column	Point	Value
I	25	±5	343
I	23	norm	805
II	23	all	divisible by 3
II	13	±4	−135
IV	21	+3	+79564
IV	19	±3	±7372
VI	15	±2	−44
VI	5	+2	+2

[a] Refers to tables given in (*1*).

Thus, we may write

$$\mathbf{y} = \mathbf{D} + \epsilon \quad (3)$$

where **D** is the $(n \times 1)$ vector of exact values that are searched.

Let us suppose now that

$$\mathbf{D} = \mathbf{X}\Theta \quad (4)$$

where **X** is a $(n \times k)$ matrix of known coefficients, with $n > k$, and Θ a $(k \times 1)$ vector of parameters. This means that the n true $\mathbf{D}_i$ values can be expressed by means of a polynomial of degree $(k - 1)$.

Thus from Relations 3 and 4, we have

$$\mathbf{y} = \mathbf{X}\Theta + \epsilon \quad (5)$$

which allows the determination of Θ and thus of **D** by Equation 4.

Then the least square method of estimation of this vector of parameters requires that we minimize the scalar sum of squares.

Table II. Corrected Version of Table V in Reference 1

CONVOLUTING INTEGERS POLYNOMIAL DEGREE = 6 DERIVATIVE ORDER = 1

POINTS	25	23	21	19	17	15	13	11	9	7
−12	−8322182.									
−11	6024183.	−400653.								
−10	9604353.	359157.	−15033066.							
−9	6671883.	489687.	16649358.	−255102.						
−8	544668.	265164.	19052988.	349928.	−14404.					
−7	−6301491.	−106911.	6402438.	322378.	24661.	−78351.				
−6	−12139321.	−478349.	−10949942.	9473.	16679.	169819.	−9647.			
−5	−15896511.	−752859.	−26040033.	−348823.	−8671.	65229.	27093.	−573.		
−4	−17062146.	−878634.	−34807914.	−604484.	−32306.	−130506.	−12.	2166.	−254.	
−3	−15593141.	−840937.	−35613829.	−686099.	−43973.	−266401.	−33511.	−1249.	1381.	−1.
−2	−11820675.	−654687.	−28754154.	−583549.	−40483.	−279975.	−45741.	−3774.	−2269.	9.
−1	−6356625.	−357045.	−15977364.	−332684.	−23945.	−175125.	−31380.	−3084.	−2879.	−45.
0	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1	6356625.	357045.	15977364.	332684.	23945.	175125.	31380.	3084.	2879.	45.
2	11820675.	654687.	28754154.	583549.	40483.	279975.	45741.	3774.	2269.	−9.
3	15593141.	840937.	35613829.	686099.	43973.	266401.	33511.	1249.	−1381.	1.
4	17062146.	878634.	34807914.	604484.	32306.	130506.	12.	−2166.	254.	
5	15896511.	752859.	26040033.	348823.	8671.	−65229.	−27093.	573.		
6	12139321.	478349.	10949942.	−9473.	−16679.	−169819.	9647.			
7	6301491.	106911.	−6402438.	−322378.	−24661.	78351.				
8	−544668.	−265164.	−19052988.	−349928.	14404.					
9	−6671883.	−489687.	−16649358.	255102.						
10	−9604353.	−359157.	15033066.							
11	−6024183.	400653.								
12	8322182.									
NORM	429214500.	18747300.	637408200.	9806280.	503880.	2519400.	291720.	17160.	8580.	60.

Table III. Corrected Version of Table VII in Reference 1

CONVOLUTING INTEGERS POLYNOMIAL DEGREE = 4 DERIVATIVE ORDER = 2

POINTS	25	23	21	19	17	15	13	11	9	7
−12	−143198.									
−11	10373.	−115577.								
−10	99385.	20615.	−12597.							
−9	137803.	93993.	3876.	−32028.						
−8	138262.	119510.	11934.	15028.	−2132.					
−7	112067.	110545.	13804.	35148.	1443.	−31031.				
−6	69193.	78903.	11451.	36357.	2691.	29601.	−2211.			
−5	18285.	34815.	6578.	25610.	2405.	44495.	2970.	−90.		
−4	−33342.	−13062.	626.	8792.	1256.	31856.	3504.	174.	−126.	
−3	−79703.	−57645.	−5226.	−9282.	−207.	6579.	1614.	146.	371.	−13.
−2	−116143.	−93425.	−10061.	−24867.	−1557.	−19751.	−971.	1.	151.	67.
−1	−139337.	−116467.	−13224.	−35288.	−2489.	−38859.	−3016.	−136.	−211.	−19.
0	−147290.	−124410.	−14322.	−38940.	−2820.	−45780.	−3780.	−190.	−370.	−70.
1	−139337.	−116467.	−13224.	−35288.	−2489.	−38859.	−3016.	−136.	−211.	−19.
2	−116143.	−93425.	−10061.	−24867.	−1557.	−19751.	−971.	1.	151.	67.
3	−79703.	−57645.	−5226.	−9282.	−207.	6579.	1614.	146.	371.	−13.
4	−33342.	−13062.	626.	8792.	1256.	31856.	3504.	174.	−126.	
5	18285.	34815.	6578.	25610.	2405.	44495.	2970.	−90.		
6	69193.	78903.	11451.	36357.	2691.	29601.	−2211.			
7	112067.	110545.	13804.	35148.	1443.	−31031.				
8	138262.	119510.	11934.	15028.	−2132.					
9	137803.	93993.	3876.	−32028.						
10	99385.	20615.	−12597.							
11	10373.	−115577.								
12	−143198.									
NORM	17168580.	11248380.	980628.	1961256.	100776.	1108536.	58344.	1716.	1716.	132.

$$\mathbf{S} = (\mathbf{y} - \mathbf{X}\Theta)'(\mathbf{y} - \mathbf{X}\Theta) \tag{6}$$

for variation in the components of Θ.

A necessary condition that Equation 6 be minimized is that

$$\frac{\partial S}{\partial \Theta} = 0$$

Differentiating we have

$$2\mathbf{x}'(\mathbf{y} - \mathbf{X}\Theta) = 0$$

which gives as least square estimator of Θ the vector

$$\Theta = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{y} \tag{7}$$

(where we assume that the matrix $(\mathbf{X}'\mathbf{X})$ is non-singular and can therefore be inverted).

Let us now consider an interval $(-m, +m)$ of the abscissas of the observations with $(2m + 1)$ unit subdivisions, we now draw our attention to a least square estimation of the central value of vector **D**, so that we may rewrite Equation 4 as

$$\mathbf{D}_{s,0} = \hat{\Theta}_1 \tag{8}$$

where $\hat{\Theta}_1$ is the first component of the estimate $\hat{\Theta}$ (the indice s is used to distinguish the true value **D** from its estimate)

Table IV. Corrected Version of Table IX in Reference 1

CONVOLUTING INTEGERS — POLYNOMIAL DEGREE = 6 — DERIVATIVE ORDER = 3

POINTS	25	23	21	19	17	15	13	11	9	7
−12	284372.									
−11	−144463.	49115.								
−10	−293128.	−32224.	748068.							
−9	−266403.	−55233.	−625974.	15810.						
−8	−146408.	−43928.	−908004.	−16796.	1144.					
−7	5131.	−16583.	−598094.	−20342.	−1547.	8281.				
−6	144616.	13632.	−62644.	−9818.	−1508.	−14404.	1430.			
−5	244311.	38013.	448909.	4329.	−351.	−10379.	−3267.	129.		
−4	290076.	51684.	787382.	15546.	876.	1916.	−1374.	−402.	100.	
−3	279101.	52959.	887137.	20525.	1595.	11671.	1633.	−11.	−457.	1.
−2	217640.	42704.	749372.	18554.	1604.	14180.	3050.	340.	256.	−8.
−1	118745.	23699.	425412.	10868.	983.	9315.	2252.	316.	459.	13.
0	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1	−118745.	−23699.	−425412.	−10868.	−983.	−9315.	−2252.	−316.	−459.	−13.
2	−217640.	−42704.	−749372.	−18554.	−1604.	−14180.	−3050.	−340.	−256.	8.
3	−279101.	−52959.	−887137.	−20525.	−1595.	−11671.	−1633.	11.	457.	−1.
4	−290076.	−51684.	−787382.	−15546.	−876.	−1916.	1374.	402.	−100.	
5	−244311.	−38013.	−448909.	−4329.	351.	10379.	3267.	−129.		
6	−144616.	−13632.	62644.	9818.	1508.	14404.	−1430.			
7	−5131.	16583.	598094.	20342.	1547.	−8281.				
8	146408.	43928.	908004.	16796.	−1144.					
9	266403.	55233.	625974.	−15810.						
10	293128.	32224.	−748068.							
11	144463.	−49115.								
12	−284372.									
NORM	57228600.	7498920.	84987760.	1307504.	67184.	335920.	38896.	2288.	1144.	8.

Table V. Corrected Version of Table X in Reference 1

CONVOLUTING INTEGERS — POLYNOMIAL DEGREE = 4 — DERIVATIVE ORDER = 4

POINTS	25	23	21	19	17	15	13	11	9	7
−12	1518.									
−11	253.	1463.								
−10	−517.	133.	969.							
−9	−897.	−627.	0.	612.						
−8	−982.	−950.	−510.	−68.	52.					
−7	−857.	−955.	−680.	−388.	−13.	1001.				
−6	−597.	−747.	−615.	−453.	−39.	−429.	99.			
−5	−267.	−417.	−406.	−354.	−39.	−869.	−66.	6.		
−4	78.	−42.	−130	−168.	−24.	−704.	−96.	−6.	14.	
−3	393.	315.	150.	42.	−3.	−249.	−54.	−6.	−21.	3.
−2	643.	605.	385.	227.	17.	251.	11.	−1.	−11.	−7.
−1	803.	793.	540.	352.	31.	621.	64.	4.	9.	1.
0	858.	858.	594.	396.	36.	756.	84.	6.	18.	6.
1	803.	793.	540.	352.	31.	621.	64.	4.	9.	1.
2	643.	605.	385.	227.	17.	251.	11.	−1.	−11.	−7.
3	393.	315.	150.	42.	−3.	−249.	−54.	−6.	−21.	3.
4	78.	−42.	−130.	−168.	−24.	−704.	−96.	−6.	14.	
5	−267.	−417.	−406.	−354.	−39.	−869.	−66.	6.		
6	−597.	−747.	−615.	−453.	−39.	−429.	99.			
7	−857.	−955.	−680.	−388.	−13.	1001.				
8	−982.	−950.	−510.	−68.	52.					
9	−897.	−627.	0.	612.						
10	−517.	133.	969.							
11	253.	1463.								
12	1518.									
NORM	1430715.	937365.	408595.	163438.	8398.	92378.	4862.	143.	143.	11.

since in that case the matrix **X** can be written

$$\mathbf{X}_{(2m+1,k)} = \begin{vmatrix} 1 & (-m)^1 \ldots & \ldots (-m)^{k-1} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & 0 \ldots & \ldots 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & (+m)^1 \ldots & \ldots (+m)^{k-1} \end{vmatrix} \qquad (9)$$

In a similar way, we deduce from Equation 4 that, in that case

$$\left(\frac{\partial^P \mathbf{D}s}{\partial x^P}\right)_{x=0} = p!\hat{\Theta}_{p+1} \qquad p = 1, \ldots, k-1 \qquad (10)$$

Thus the p-derivative of the adjusted vector **D** at the central point is equal to the product of $p!$ by the $(p + 1)$ component of the estimated vector $\hat{\Theta}$.

Now from Relation 7 it appears that $\hat{\Theta}$ is the product of the observation vector **Y** by a $(k \times n)$ matrix T

$$\mathrm{T} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}' \qquad (11)$$

of known coefficients.

Table VI. Corrected Version of Table XI in Reference 1

CONVOLUTING INTEGERS			POLYNOMIAL DEGREE = 6				DERIVATIVE ORDER = 5			
POINTS	25	23	21	19	17	15	13	11	9	7
−12	−1012.									
−11	253.	−209.								
−10	748.	76.	−3876.							
−9	753.	171.	1938.	−102.						
−8	488.	152.	3468.	68.	−104.					
−7	119.	77.	2618.	98.	91.	−1001.				
−6	−236.	−12.	788.	58.	104.	1144.	−22.			
−5	−501.	−87.	−1063.	−3.	39.	979.	33.	−3.		
−4	−636.	−132.	−2354.	−54.	−36.	44.	18.	6.	−4.	
−3	−631.	−141.	−2819.	−79.	−83.	−751.	−11.	1.	11.	−1.
−2	−500.	−116.	−2444.	−74.	−88.	−1000.	−26.	−4.	−4.	4.
−1	−275.	−65.	−1404.	−44.	−55.	−675	−20.	−4.	−9.	−5.
0	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1	275.	65.	1404.	44.	55.	675.	20.	4.	9.	5.
2	500.	116.	2444.	74.	88.	1000.	26.	4.	4.	−4.
3	631.	141.	2819.	79.	83.	751.	11.	−1.	−11.	1.
4	636.	132.	2354.	54.	36.	−44.	−18.	−6.	4.	
5	501.	87.	1063.	3.	−39.	−979.	−33.	3.		
6	236.	12.	−788.	−58.	−104.	−1144.	22.			
7	−119.	−77.	−2618.	−98.	−91.	1001.				
8	−488.	−152.	−3468.	−68.	104.					
9	−753.	−171.	−1938.	102.						
10	−748.	−76.	3876.							
11	−253.	209.								
12	1012.									
NORM	1300650.	170430.	1931540.	29716.	16796.	83980.	884.	52.	26.	2.

If we give to n the $(2m + 1)$ values that were considered above and for a fitted polynomial of degree $(k - 1)$, we may tabulate the values of the elements of these matrix T.

Therefore, the convolution of the first line element of this matrix T with the observation data will give the least square smoothed value of the central point of the data vector whereas the convolution of the elements of the $(p + 1)$ line, multiplied by $(p!)$, with data, will give the p-derivative at this point.

In fact, if we refer to Savitzky and Golay formalism, we have

$$\sum_{k=0}^{n} b_{n,k} S_{r+k} = F_k$$

where

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

$$F_k = \sum_{i=-m}^{i=+m} i^k y_i$$

where the $b_{n,k}$ are the coefficients of the fitted polynomial of degree n, i denotes the abscissas of the $(2m + 1)$ considered points and is therefore an integer varying from $-m$ to $+m$. Using our notation, these expressions are equivalent to

$$(\mathbf{x}'\mathbf{x})\hat{\Theta} = \mathbf{X}'\mathbf{Y}$$

More details about this method of calculation and its generalization will be reported in a following publication.

The corrected values are presented in Table I. Also, Tables II, III, IV, V, and VI are corrected versions of Tables V, VII, IX, X, and XI given by Savitzky and Golay (*1*). They either contain more numerous errors or are systematically false.

ACKNOWLEDGMENT

We thank P. T. Dagnelie and the staff of his department for providing computing facilities and help in programming.

JEAN STEINIER
YVES TERMONIA
JULES DELTOUR

Laboratoire Central du Camira et
Department de Physique et Chimie Physique
Faculté des Sciences Agronomiques
5800 Gembloux, Belgium

RECEIVED for review October 4, 1971. Accepted April 7, 1972. This research program was carried out partially under the sponsorship of the "Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture" of Belgium.

"Trading Rules" in Infrared Fourier Transform Spectroscopy

SIR: In their recent article, Pickett and Strauss (*1*) discussed various factors affecting the signal-to-noise ratio (S/N) in IR and NMR Fourier Transform Spectroscopy (FTS) (FTS is a registered trademark of Digilab Inc.). They concluded that for optimum analytical sensitivity, the resolution should be chosen such that the half-width of the bands being measured is only slightly greater than the resolution used in the measurement. This criterion has, of course, been used in dispersion spectroscopy for many years. The actual quantitative relationships between S/N, resolution ($\Delta\nu$), and measurement time (T) in IR Fourier Transform Spectroscopy, commonly referred to as "trading rules," (*2*)

(1) H. M. Pickett and H. L. Strauss, ANAL. CHEM., **44**, 265 (1972).

(2) W. J. Potts and A. Lee Smith, *Appl. Opt.*, **6**, 257 (1967).