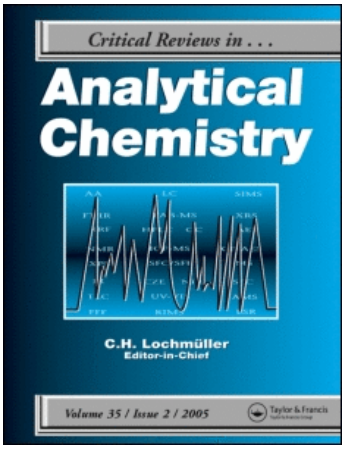


Critical Reviews in Analytical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713400837

Chromatographic Data Systems: A Critical Review

Andrew N. Papas[a]
[a] Analytical Branch U.S. Food and Drug Administration, Winchester, Massachusetts

To cite this Article Papas, Andrew N.(1989) 'Chromatographic Data Systems: A Critical Review', Critical Reviews in Analytical Chemistry, 20: 6, 359 — 404
To link to this Article: DOI: 10.1080/10408348908050072
URL: http://dx.doi.org/10.1080/10408348908050072

CHROMATOGRAPHIC DATA SYSTEMS: A CRITICAL REVIEW

Author: **Andrew N. Papas**
Analytical Branch
U.S. Food and Drug Administration
Winchester, Massachusetts

Referee: Stephen L. Morgan
Department of Chemistry
University of South Carolina
Columbia, South Carolina

I. INTRODUCTION

A. Perspective and Goal

This review article spans over 30 years of chromatography advancement and development. Coincident with the evolution of chromatography was that of laboratory data automation. The first commercial gas chromatograph appeared in 1955,[1] whereas the first commercial digital integrator appeared in 1961.[2] Several data systems emerged in the late 1960s[3-6] that led to the arrival of the first commercial computer-based electronic integrator in 1972.[7] Two major advancements were most influential to the development of data automation; initially there was the development of solid-state operational amplifiers that led to analog peak integrators. Prior to this time, manual peak integration methods were used. Then the advancement of integrated circuits allowed digital logic to be introduced. Today, all commercial and noncommercial laboratory integrators are sophisticated computer-based systems. This article generally follows this chronological development.

This review evaluates the development and validation of a completely automated, accurate, and precise chromatography peak-area integrator that works reliably under *any* chromatographic conditions. All assessments are made with this goal.

Electronic integrator development has spanned many disciplines, including chromatography theory, probability and statistics, analog circuit design, digital logic and design, signal processing, and computer science. Each of these individual aspects can be independently optimized, yet are interdependent in the final product. The ultimate test of these devices is how well it quantitates real data. Therefore, both pertinent components and the overall assembled system are reviewed.

Methods for resolving overlapped peaks, noise characterization and filtering, peak detection schemes, evaluation techniques, and results as well as baseline construction methods are reviewed individually; all are involved in either converting raw data into peak areas or assessing performance. Peripheral functions generally found on present-day, computer-based systems such as networking, off-line calibration and calculations, interfacing to data bases or laboratory information management systems (LIMS), etc. are not considered. Furthermore, where the distinct differences between an integrator and a data system are based only on these auxiliary functions and not on different peak-detection and quantitation algorithms, no distinctions are made here. The two terms are used synonymously.

The alogrithms and hardware of data systems of today were assessed and developed many years ago. To understand what is available today and what might become available tomorrow, one must look at the development of data processing over time. Much more has been written on computer automation and processing of chromatographic data over 30 years than is referenced here. It would neither be possible nor necessary to cite all publications. Included are representative papers in the topic areas, whereas all major advancements are necessarily cited. Present-day systems are a product of compromise of many competing factors. Future

systems will be constrained by a different set of compromising factors with these same principles. Thus, with an eye toward the future, these general topics are reviewed.

B. Outline

Peak area is generally considered to be a more "fundamental" measurement than peak height. Chromatography theory concerns the diffusion or band-broadening of a mass as it passes through the chromatographic column. The overall result is a peak envelope or a distribution vs. time profile. Its sum represents the mass of analyte[8-10] measured by its area. With concentration sensitive detectors, peak areas have been shown to be independent of chromatographic drift or compositional changes, but adversely affected by flow rate changes[11,12] in comparison to peak height measurements. Instrumentally, peak areas have been shown to be more accurate overall yet less precise,[13-19] where the precision is predominantly limited by the uncertainty in determining either the baseline or peak end.[16-19] Therefore, since peak areas are more sensitive to measurement errors and are generally more desirable for quantitative analysis, they are the measurement assessed in this review.

Two fundamental measurements restrict the accuracy and precision of peak-area determinations: correct baseline construction and the resolution of overlapped peaks. These problems are neither peculiar to chromatography nor adequately satisfied in other disciplines. Each is addressed differently. Baseline construction is discussed throughout this paper for each peak-sensing device and is not presented separately. However, peak-resolution techniques are reviewed individually, even though most, if not all, integrators today utilize the same perpendicular drop method. The reason is twofold: the perpendicular drop is not accurate in resolving fused peaks; yet, it is the most reliable method. Its use is a compromise of accuracy for reliability and future commercial laboratory peak-resolution techniques will involve one or several of these other resolution techniques, possibly in combination.

When investigated, noise has been shown to adversely affect almost all algorithms involved in either peak detection, peak resolution, or peak quantitation. Several methods for noise reduction have been proposed, yet can only be applied within certain restraints. Proposed peak-sharpening algorithms that enhance resolution of unresolved peaks are based on filtering techniques. Therefore, this important topic is discussed separately, focusing primarily on digital (computer) logic systems.

II. CHARACTERIZATION AND RESOLUTION OF OVERLAPPING PEAKS

Chromatographers have developed efficient columns for the separation of complex mixtures. Yet many times additional efficiency reveals partially resolved coeluting bands not seen with older-generation columns or instrumentation. With complex sample matrices such as biological, agricultural, or natural product separations, it may not be possible to avoid overlapping peaks. In many cases either pure standards are not available or the standard "background" does not match that of the analytes. In these cases the accuracy of the peak-resolution or deconvolution software is very critical. This problem becomes most acute when one is looking at peaks that approach the limits of detection and are not baseline resolved. Even though it is recommended that all quantitative analyses involve baseline-resolved peaks to avoid problems,[15] this is not always possible. These cases rely on their intelligent integrator/data systems to quantitate their analyte reliably.

Compared to algorithms that determine baseline levels or determine a start and end of a single peak, peak-resolution software is the most difficult to develop and optimize. It represents the most time-consuming effort of the development of any peak integrator sometimes takes years to debug. Many chromatographers have developed their own peak integrator/data reduction systems.[20-30] In many cases, this is justified on the grounds of avoiding a costly commercial system. However, many systems either avoid peak resolution altogether and

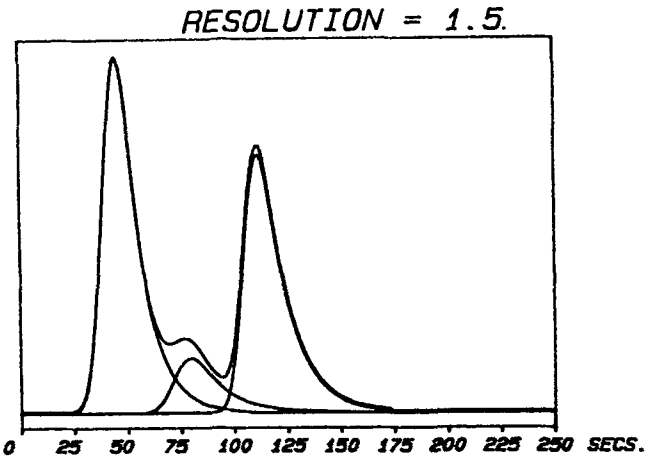


FIGURE 1. Overlapped and skewed peaks. Resoltuion = 1.5; tau/sigma = 2.2

offer no capability to resolve peaks,[20-23] or they admit their algorithms are not optimal or are limited.[24-30] In these cases it is not clear what is done with unresolved peaks. With modern chromatography, adequate peak-resolution routines are mandatory.

Even the measure of resolution is not without debate. Textbook resolution calculations based on the characteristics of ideal Gaussian peaks have discrepancies that become proportionally larger as the peaks become tailed.[31-33] Kirkland et al.[33] pointed out the ambiguity of resolution calculations when the square root of the second moment was used for peak width. He suggests that new calculations be developed that take into account the third moment or skew. Yet most authors presenting work on resolving tailed peaks measure the resolution with the Gaussian calculations. To avoid additional confusion, this paper cites the calculation method used.

When encountering two or more partially overlapped peaks (Figure 1), the completely automatic, accurate, and reliable integrator might be expected to possess the following ideal characteristics:

1. Be able to determine the peak purity and number of components present
2. Make no assumptions about peak shape
3. If based on a model, be able to provide confidence levels to the chromatographer
4. Be able to adequately resolve and quantitate the components, regardless of the degree of overlap or resolution
5. Be insensitive to noise
6. Be unaffected by the amount of peak tailing
7. Assume unlimited computer speed, memory, and power
8. Be unaffected by the weight percent of components present
9. Be precise, accurate, and reliable
10. Be based on single-channel data
11. Easily handle nonlinear and shifting baselines
12. Require no prior knowledge of the components present nor require independent standard chromatograms
13. Work either in real-time or within a reasonable time frame

No such device exists today, and resolution of overlapping peaks still is a very active area of research. Here we investigate several approaches to resolve overlapping chromato-

graphic peaks. Each method has its own merits and limited applicability; yet each becomes restricted by some of the above-mentioned items.

A. Moment Analysis and Peak Models

An infinitely variable waveform requires an infinite number of parameters to describe it precisely. Chromatography can be considered the process(es) to convert a thin sample slug at the beginning of the column into an elution profile whose shape is governed by physical and chemical interactions with the column and mobile phase. One goal of chromatography theory is to observe and characterize the resultant waveform to better understand the processess occurring and their sphere of influence. This leads to more general mathematical models that adequately describe this elution profile with a finite set of parameters. From the data-processing standpoint, peaks need not be described point-by-point, but instead by a reduced set of parameters. General models lead to simplifications in peak-resolution techniques. However, the danger is that real data may not fit the assumed model adequately, and the level of error is related to its lack of fit. Thus, a general model for typically encountered data helps describe the chromatographic processes and leads to simplification in data analysis.

Chromatography theory has developed many independent approaches to describe the physical interactions. Based on the theoretical plate model, an elution profile can be described as a binomial distribution when the plate count is low, whereas it approaches a Gaussian profile when the plate count is above 100.[8] Sorption-desorption theory predicts chromatographic profiles to have the form of a Bessel function and exponential term.[9] No mathematical model acceptable to all theories has been found. More complex and asymmetric functions require more parameters to describe the elution profile and require more calculations to quantify chromatographic figures-of-merit. For ease and simplicity, most basic chromatography relationships such as resolution and plate height used today assume the ideal case of a Gaussian profile.

By analyzing peak profiles of real data with moment analysis, one can test a proposed model for validity. The moments calculated are then compared to the expected values of the moments derived from the particular model using calculus. The differences observed reflect the validity of the model. Once the validity of the model has been demonstrated, its use leads to simplifications. It is particularly important to use a model valid for most chromatographic peaks when simulated chromatography is used to evaluate the quantitation of an integrator. A validated model is also crucial when deconvoluting unresolved peaks by curve-fitting techniques.

Moment analysis has few restraints for its application and is the most general technique available to completely describe a peak. No assumptions are made about peak shape, and all peak profiles can be treated similarly. The zero through fourth moments are given in Figure 2. From a chromatographic viewpoint the zeroth moment represents the peak area, the first moment is the center of mass and approximates the retention time, the second moment represents the variance or standard deviation squared, the third central moment represents the skewness or asymmetry (related to the peak-tailing factor), and the fourth central moment is the excess. These chromatographic equivalences are independent of the mathematical model proposed. For some functions these moments are expressed in convenient terms. For example, for an exponentially modified Gaussian (EMG) probability density function (see Figure 3), the first moment is tg + tau; the second is $sigma^2 + tau^2$; the third is $2*tau^3$; and the fourth is $3*sigma^4 + 6*sigma^2*tau^2 + tau^4$.

Imprecision is the one drawback to validating models using numerical moment analysis. Moment analysis has been shown to be sensitive to noise,[31,34,35] peak-start and -stop assignment,[31,33] baseline drift,[36] correct sampling rate,[34] and nonlinearity of the distribution isotherms.[37] These effects are particularly amplified for the higher moments.

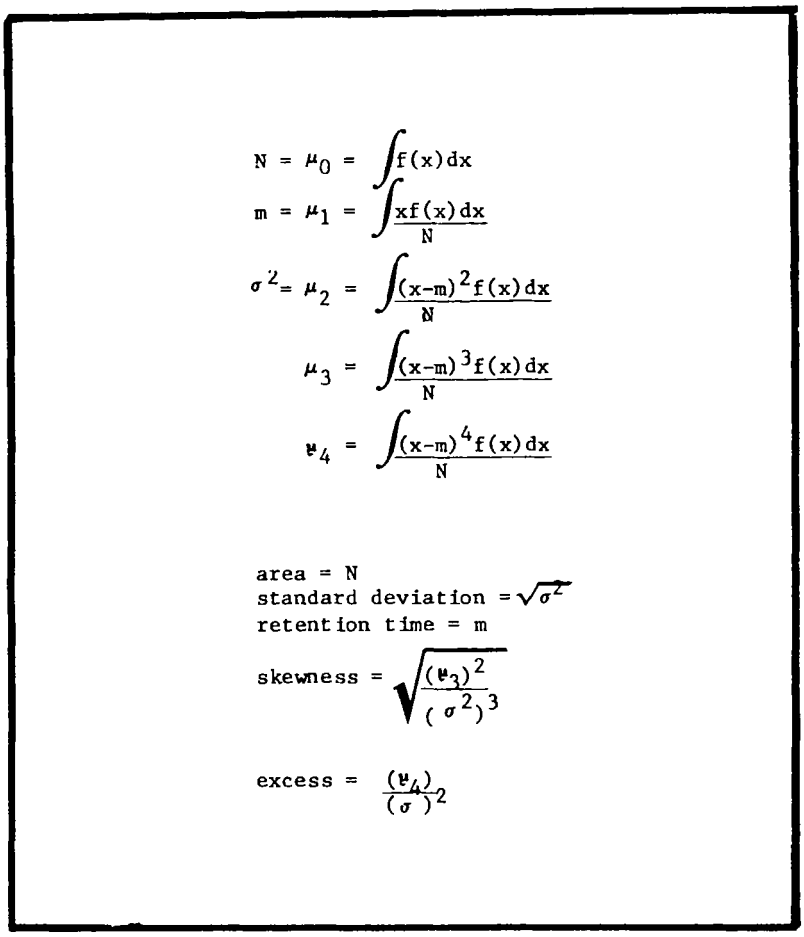


FIGURE 2. Peak moments. Zeroth moment is the mass or peak area. First moment is the center of mass or retention time. The second moment is the variance or standard deviation squared. The third moment is the skew. The fourth moment is the excess.

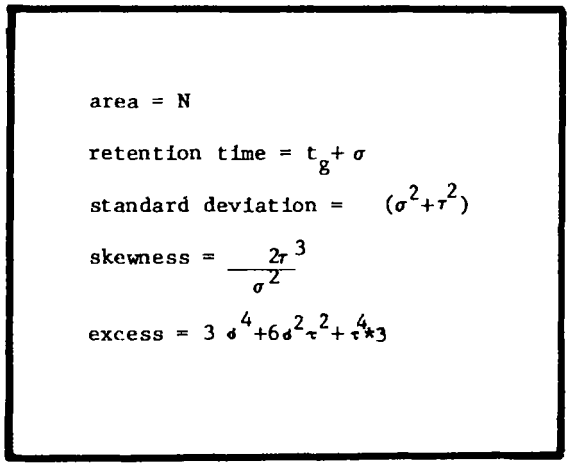


FIGURE 3. Moments for the EMG, expressed in tau, sigma, and tg - - - the retention time for the nonskewed Gaussian peak.

Despite these imprecisions, numerical moment analysis has been extended beyond peak modeling and has been used to discern the purity of peaks,[38] where the skew and excess of simulated biGaussian and Poisson overlapped peaks show relationships different from those obtained from single peaks. When applied to real data, though, the results are somewhat disappointing as ambiguities arise. This is only one of several approaches to discern peak purity with single-channel data.

The general idea that third and fourth peak-moment ratios contain information about peak overlap and purity is not new. Yet, in the past the utility of this approach was limited by the means of obtaining the moments. Jung et al.[39] used particular characteristics of the EMG profile to derive the moments. This was shown to be susceptible to noise and become unreliable as the peak deviates from the EMG model. Grushka[40] compared moment analysis to the second derivative method and found them almost equivalently susceptible to noise and equivalently imprecise with large peak-mass ratios. A simpler method based on points on inflection relationships by Grubner[41] only works in the ideal case. Recently presented graphic relationships derived analytically for the EMG should facilitate matters.[42-46] In fact, calculated moments are available on some integrators or can be easily programmed.[47] At present, chromatographers may make use of these EMG moments for many things other than peak purity. However, moments are not generally used in peak-resolution or peak-detection algorithms.

As mentioned above, chromatographic peaks approximate Gaussian profiles without instrumental distortion. However, it has been shown many times that any number of internal and external extracolumn processes lead to peak asymmetry.[33,48-50] Some feel that as a result of too many undefined causes for tailing, no unique analytic functional form can be used for such peaks.[51,52] Yet for the above-mentioned reasons, a representative functional form has been pursued vigorously. Due to instrumental distortions of the ideal Gaussian profile, true chromatographic peaks are better represented by a profile that is asymmetric.

Models that have been proposed and applied to real data include: a combination Gaussian/triangle/exponential decay;[34] a combination Cauchy and Gaussian;[53,54] a combination Gaussian/hyperbolic/exponential decay;[55] a biGaussian;[38,49] Poisson;[38,49] a cam-driven analog peak;[31] a binomial;[50] and a Gram-Charlier.[41] None approach the extent of theoretical and experimental work done by Pauls and Rogers,[56] Grushka,[38] Maynard and Grushka,[57] and others on demonstrating exponential decay which justify the EMG model. More recent work by Foley[44] has shown this model to be valid for over 90% of the real data analyzed. On this assumption of the EMG being a valid model, others have proposed that entire chromatograms be saved as just moments of the EMG.[58] Based on the current evidence, it would appear that the EMG is the best model proposed so far to mimic real chromatographic data.

While Foley,[44] showed over 90% of the real data tested can be described adequately by the EMG model, this still leaves 10% uncharacterized. When the EMG was fit to overlapped-capillary and packed-column GC peaks by least-squares curve fitting, it was found to be adequate only for the capillary peaks and not for the packed peaks.[59] In this case these differences may be attributed to the type of instrumental distortion found in each type of analysis. Others have demonstrated that different functional groups give rise to different peak shapes[60,61] and that peak shape for one compound changes with the concentration levels eluted.[62] Different retention mechanisms observed within one chromatogram invalidate applying one general function to describe the physical phenomena. Although recent simulation studies have developed extensive algorithms for discerning the causes on nonideal chromatography,[63-65] one general function will not be sufficient for our automatic, reliable, and accurate integrator if this is a prerequisite for its peak-resolution technique.

In summary, moment analysis can aid in providing a mathematical functional form of real data. It can also provide information on chromatographic conditions when data no longer fit this model. The methods of obtaining the moments vary from the general numerical

approximation to derived measurements based on a proposed function. They all are generally susceptible to the same errors encountered in determining peak areas. Moments can also aid in determining the peak purity when the chromatography is well behaved. It is not a general method for resolving overlapped peaks, however.

B. Geometric Peak-Area Methods

Unlike moment analysis, most geometric methods assume either a Gaussian profile or, minimally, a symmetric profile. They are easy to implement either graphically or in a computer. The five basic methods to measure peak areas graphically (Figure 4) are (1) peak-height times the width at half height; (2) a triangle through the two points of inflection; (3) peak height times the average of peak width at 85% height and 15% height; (4) planimetry; and (5) cutting, then weighing. These methods are reviewed later. The four graphic quantitation methods for overlapping peaks (Figure 5) are (1) the perpendicular drop method, (2) triangulation, (3) peak-height ratios, and (4) the tangent skim. These methods can be applied manually or with a computer.

Five major factors affect accurate resolution with geometric methods. They include resolution (however measured), the level of noise, the relative ratio of the two (or more) components, the amount of tailing, and the order of the smaller peak. For each method the relative contribution of each factor varies.

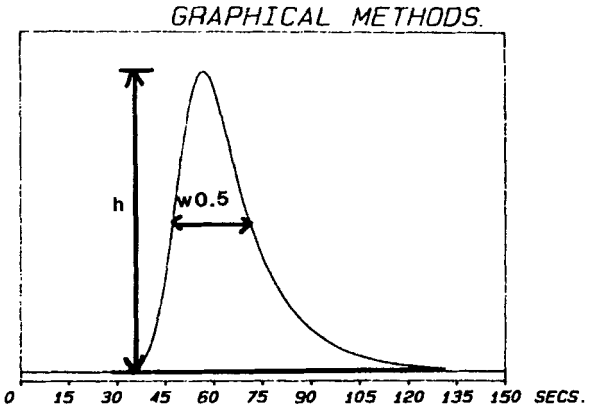
Peaks can overlap slightly, moderately, or severely. These general terms can be described in the following manner (Figure 6). Slightly overlapped peaks are indicated by two peaks that are visually distinct, and there exist two maxima and one minimum. Resolution values of 2.0 to about 1.0 (using moments) are in this range. Moderately resolved peaks involve a resolution of about 1.0 to about 0.5 or to the point where two maxima are still observed with an intervening valley point. Severely overlapped peaks include those with shoulder peaks where one maximum and one point of inflection are present down to visually indistinguishable overlapped peaks. A special case occurs when analyte peaks ride the back of a much larger-tailed solvent peak and required special attention.

Severely overlapped peaks are not considered here, as graphic methods require visual alignment for correct placement of the appropriate template. Visually indistinguishable, overlapped (coeluting) peaks are almost impossible to detect with single-channel data and no calibration standards. Multivariate techniques using either multiple channels of data or sufficient calibration are discussed below.

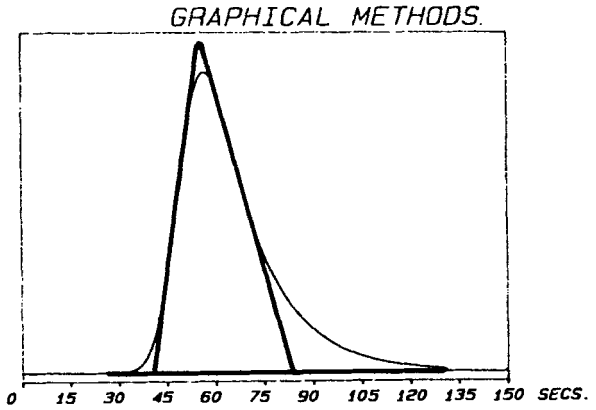
Triangulation is the simplest manual method (Figure 5a), yet is also the least accurate of all methods.[66-69] With single Gaussian peaks a triangle approximation only encompasses 93% of the true area. As overlapped peaks increasingly tail, inaccuracies become intolerable. For example, with two peaks at a height ratio of 45:1 the error with triangulation was −45%.[66] Ball et al. correctly stated that triangulation had little to offer.[71]

The perpendicular drop method (Figure 5b) makes no assumptions about the peak shape, is easily implemented within a computer, and is very reliable as only a valley point need be detected. It is currently the only method used by commercial integrators. However, it is not accurate because it cannot correctly proportion the relative contributions of each peak to the bisected area determined. The relative inaccuracies increase as peaks become more and more tailed and as the relative peak ratios vary.[17,66-69] Attempts have been made to develop correction factors to correctly apportion the peak areas.[66,67,69] However, these factors assume a Gaussian model and they themselves become inaccurate as the peaks tail. One attempt to partially correct for the relative overlapping contribution of different-sized peaks when a perpendicular drop is used is the algorithm of Landowne et al.[70] This algorithm attempts to shift the drop point toward the smaller peak to reduce the contribution of the larger peak. No comparison was given on performance, though.

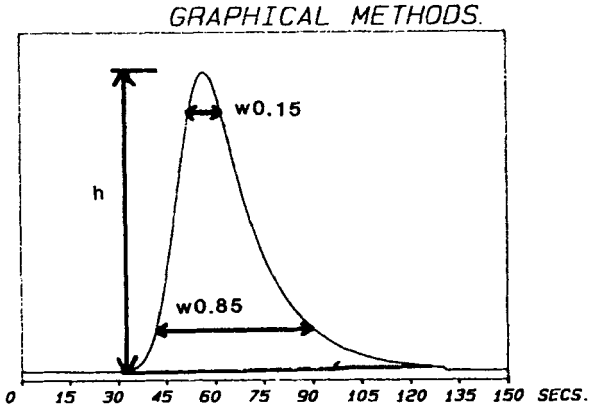
Any attempt to develop peak-area correction factors with a more appropriate model such



A



B



C

FIGURE 4. Graphical measurement of peak areas. (A) Area equals height times width at $^1/_2$ height; (B) area is approximated by triangle; (C) area = height times the average of the widths at 15 and 85% height; (D) disk integration; (E) cut, then weigh.

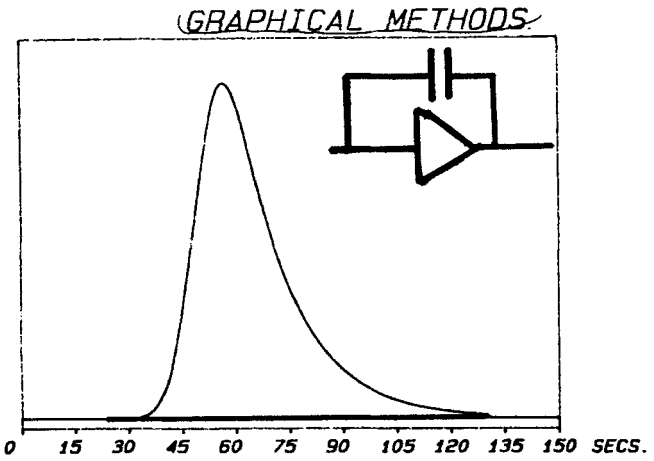


FIGURE 4D

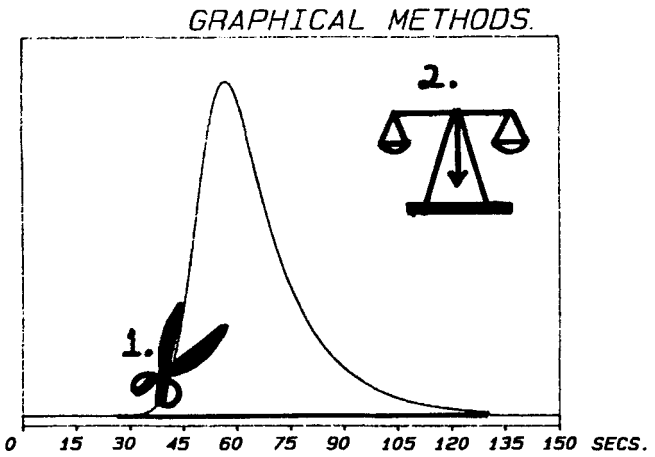
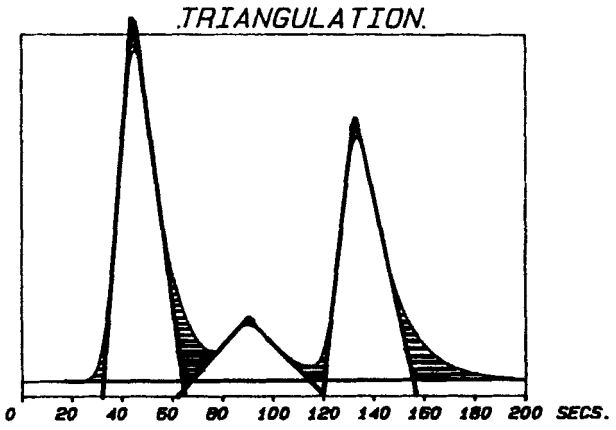


FIGURE 4E

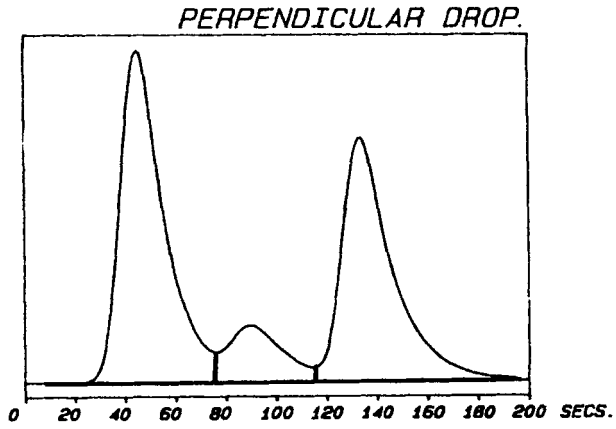
as the EMG still must struggle with the problems already mentioned with model assumptions: not all peaks fit the (EMG) model; the model changes as the concentration levels change; and the two peaks being corrected may elute by different retention mechanisms and, therefore, require different models. Even if accurate correction factors were available for any mathematical peak model at any peak resolution, there would still be the requirement that the resolution be measured accurately and the correct model be identified, shown otherwise. Therefore, no attempt is made to correct for overlapped peak tails with correction factors, presumably because these factors may introduce inaccuracies beyond that of the perpendicular drop method itself.

Given the inaccuracy of the perpendicular drop method currently used by present-day integrators, many think that external standard calibrations negate this problem. This would only be true if the standards were run with the same level of overlap at the same relative peak ratio to the interfering peak(s) (Figure 7). In most cases the standards are "cleaner" than the sample with no overlap onto other peaks. Warnings to quantitate analyte peaks only when baseline resolved are well-founded.

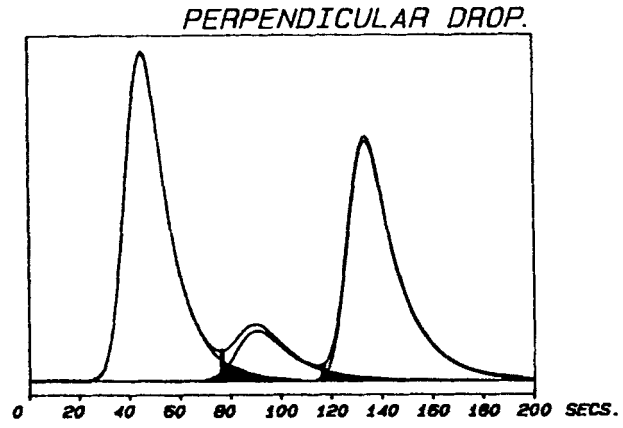
The peak-height ratio method (Figure 5c) is analogous to the peak-area perpendicular drop method. It is no more accurate than the perpendicular drop method; no compensation is



A



B1



B2

FIGURE 5. Geometric peak resolution methods. (A) Triangulation. Error is stripped region. (B1) Perpendicular drop method. (B2) Perpendicular drop method showing individual peaks and errors. (C) Peak height method. H-error is height not taken into account with overlapped peaks. (D) Tangent skimming. (E) Perpendicular drop incorrectly applied.

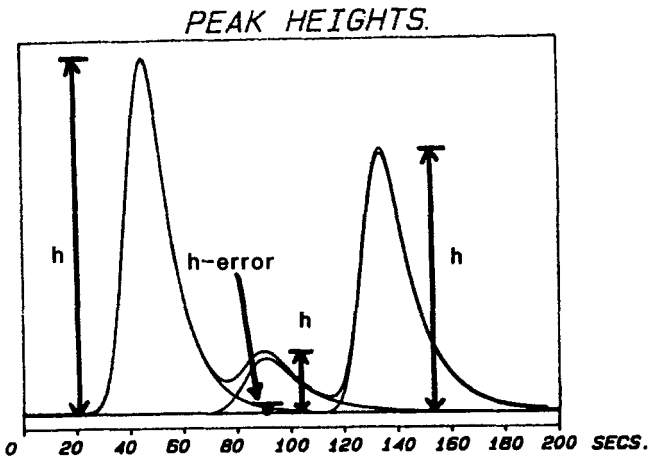


FIGURE 5C

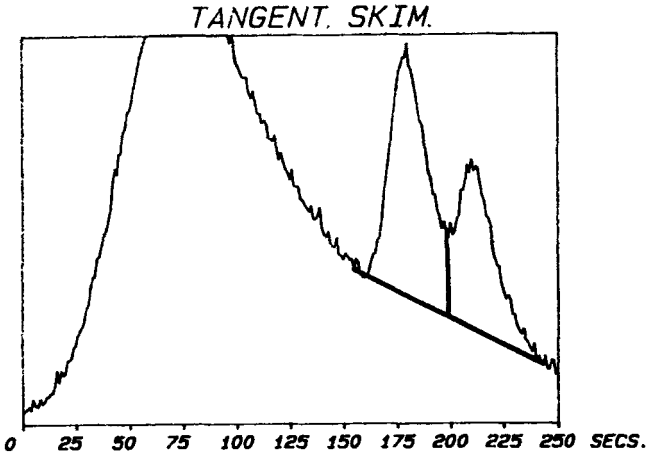


FIGURE 5D

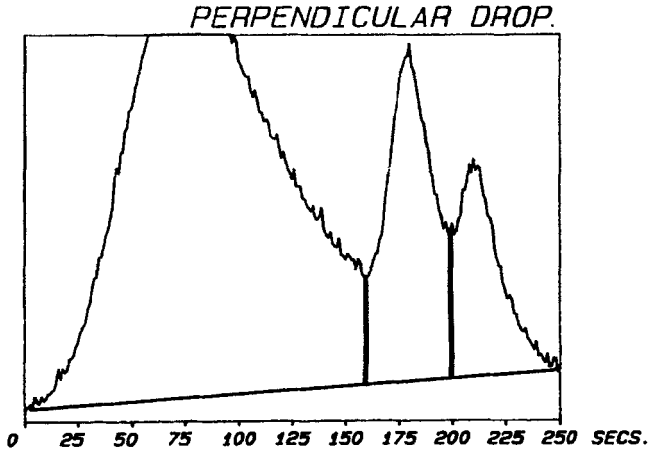
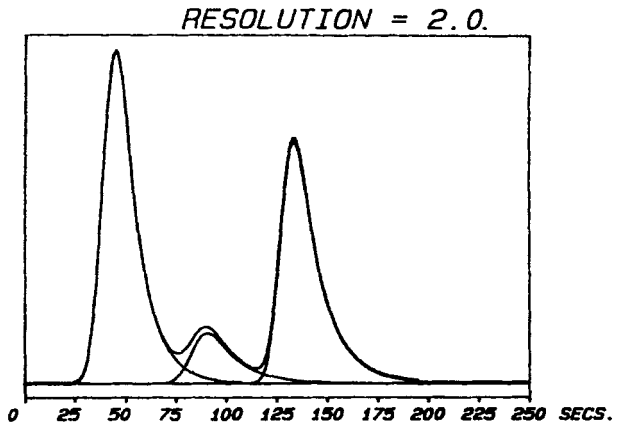
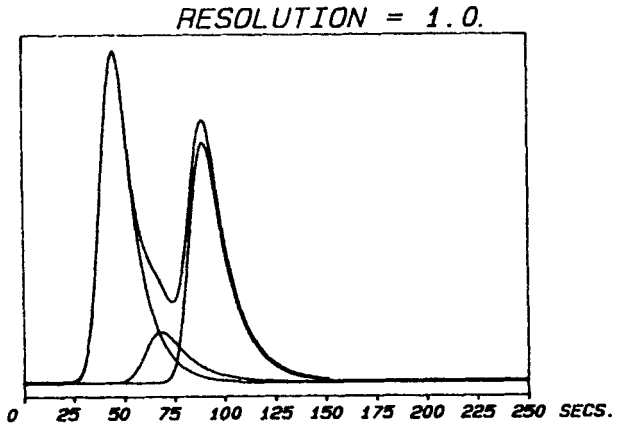


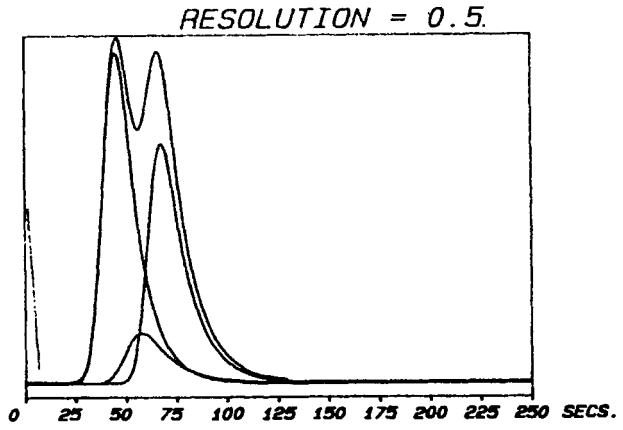
FIGURE 5E



A

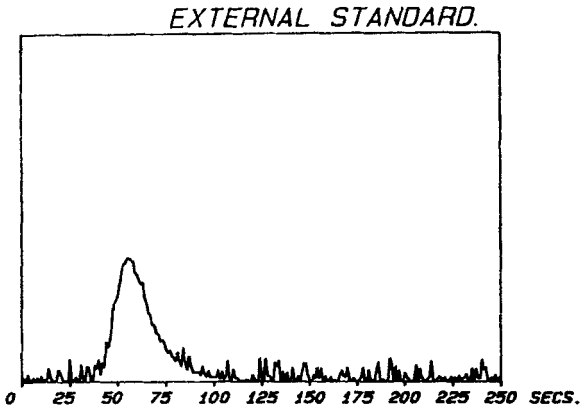


B

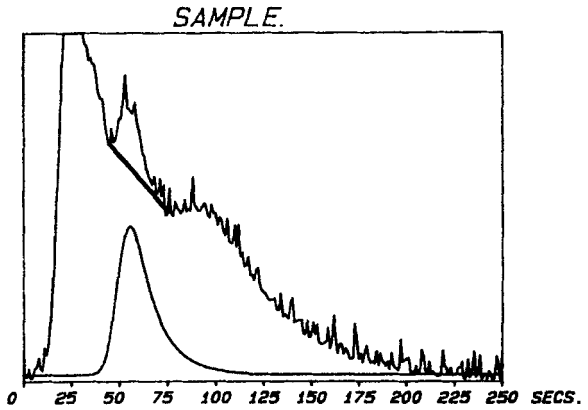


C

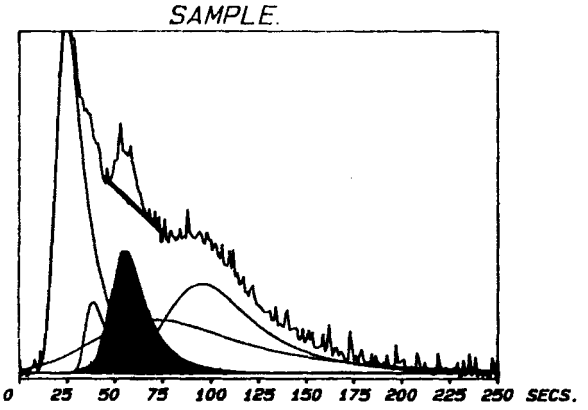
FIGURE 6. Resolution ranges. (A) Slight, resolution = 2.0. (B) Moderate, resolution 1.0. (C) Severe, resolution 0.5.



A



B



C

FIGURE 7. External standard inaccuracies. (A) Clean standard. (B) Sample peak in complex matrix; only sample peak shown along with its skimmed peak area; (C) same as (B), except all peaks that comprise the peak envelope are shown.

made for the relative contribution of each peak to the net peak height measured at the two maxima. A study by Mori[72] on developing correction factors first required precalibration with known standards at various resolution. These values were then applied to samples exhibiting the same level of overlap. This was shown to be reasonably accurate down to a resolution of 1.0 (measured by the width at half height). As resolution decayed, error increased dramatically, especially as tailing increased. This is another case where an incorrect model assumption leads to inaccuracy as the peaks drift from that model assumption.

Although peak-height correction factors could be developed for an EMG model and stored in a look-up table, this would assume accurate resolution calculations as well as accurate moment calculations. Although no work has been done in developing peak-height correction factors for the EMG, the peak-area perpendicular drop method would still be preferable as peak-height measurements are more precise and less accurate than peak areas. As chromatographic conditions change, so do the peak heights.

When the peak ratio of two overlapping, tailed peaks approach some upper limit, the perpendicular drop method is very inaccurate for the smaller peak. By dropping a perpendicular, the smaller peak is grossly overestimated, whereas the larger peak is underestimated.[67] In this case the tangent skim method is more appropriate (Figure 5d). The peak-height ratio at which to switch from the perpendicular drop method to the tangent skim method is not well defined; however, a general rule of thumb seems to be 10:1. Particularly important for the tangent skim method is the decision on when a peak ends. With proprietary algorithms, the only way of knowing the accuracy of commercial systems is to test them with real or simulated chromatography. Recent studies indicate very real and large differences exist among manufacturers, leading to large differences in peak quantitation.[74,75] Tangent skim methods are not available on some commercial data systems.[73]

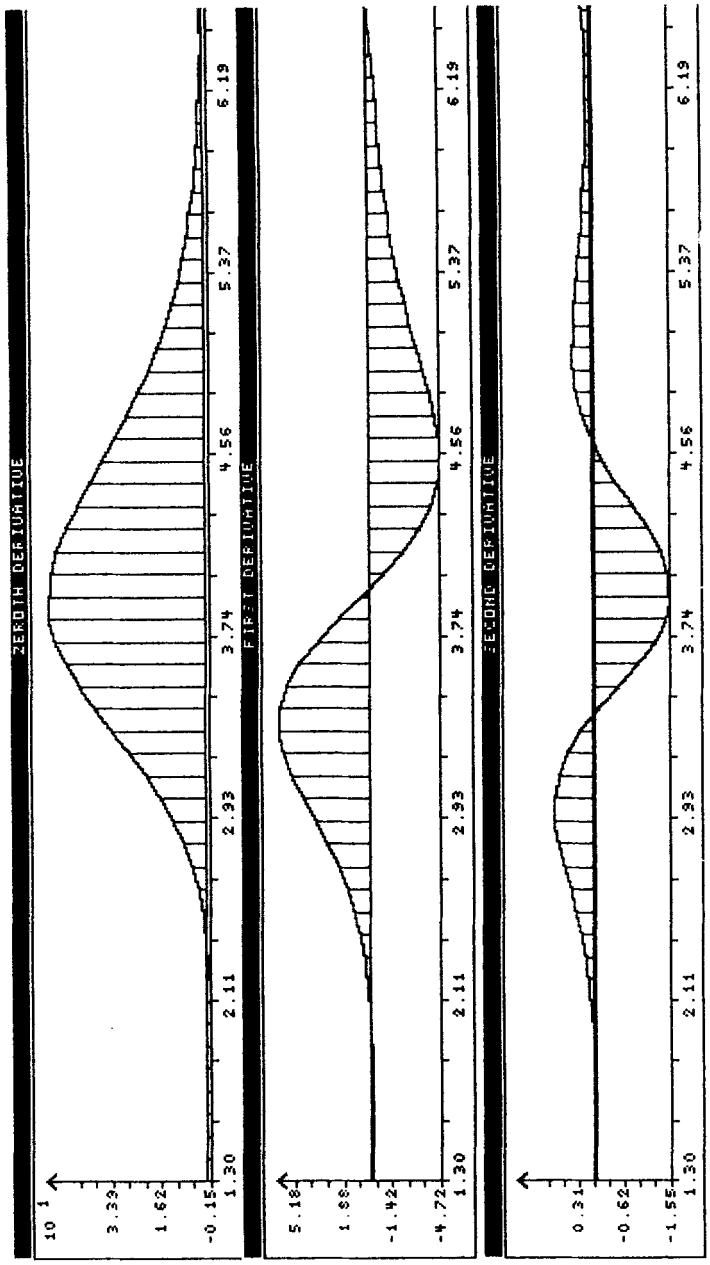
In summary, graphic methods are easy to implement either by hand or by computer. Application of these methods is very reliable with moderately overlapped peaks. However, there are inherent inaccuracies that are not adequately corrected either within the method of by using "cleaner" external standards. It is because of these inaccuracies that alternative methods are sought.

C. Derivative Peak Detection

Although some commercial algorithms claim to detect peaks without calculating derivatives explicitly,[76] virtually all peak-detection software relies on the rate of change of the signal for peak detection. This is seen in Figure 8a as the first and second derivative follow a reproducible pattern for an EMG profile. The first derivative goes positive, passes through a maximum at the inflection point, then becomes negative until the second inflection point. The zero crossing indicates the retention time, whereas the return to zero indicates the peak end. The second derivative is not needed here. This pattern is true for any peak profile; therefore, no model assumptions are necessary. For overlapped peaks (Figure 8b) a different pattern is observed. Three zero crossings are observed in the first derivative for moderately overlapped peaks, where the second derivative is used to confirm maximum and minimum. These inflection points, zero crossings, and second derivatives are easy to obtain with a computer in real-time, or with analog operational amplifiers. This method is the basic peak-detection scheme used widely in commercial integrators today.

It is generally accepted that numerical differentiation is an unstable mathematical process. White noise has been shown to reduce the signal-to-noise ratio by a factor of two for each subsequent derivative.[77] Because it is extremely susceptible to distortion and noise extremes, caution must be used in applying differentiation. Any derivative methods used for peak detection or for resolution enhancement must first reduce noise to minimal levels without distorting the signal. Noise filtering is discussed below.

As expected there are differences among methods used to obtain derivatives. Shukla and



A

FIGURE 8. Chromatograms and their derivatives. (A) Single EMG peak, tau/sigma of 2.0, with its first and second derivative. (B) Two slightly overlapped EMG peaks with their first and second derivative.

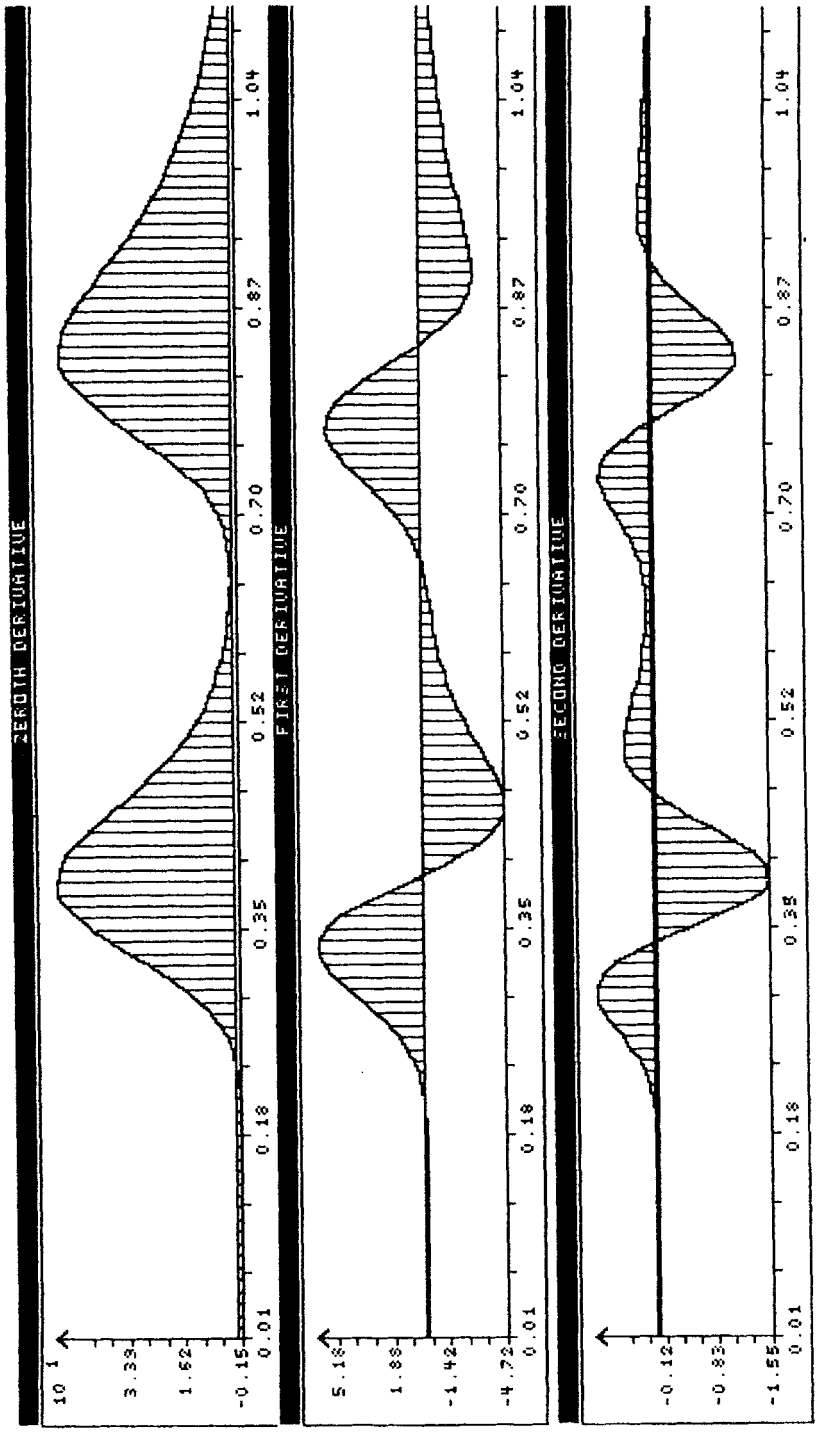


FIGURE 8B

Rusling[78] compared a five-point central difference method to a five-point Savitzky-Golay[79] quadratic polynomial method. The Savitzky-Golay method is more tolerant to noise from chromatographic and spectroscopic data, yet tends to distort the signal at higher orders of differentiation. They cautioned software users to be informed on exactly how the differentiation is done so they can be aware of possible errors and limitations.

Derivatives have been used for more than peak detection. Kambara and Satoh[80] reported on using the peak-height ratio of the derivative chromatogram graphically to quantitate poorly resolved small peaks, where the net peak height is the absolute distance from one inflection point maximum to another inflection point maximum. This method was found superior to a manual peak area by triangulation, but can only be applied to moderately resolved peaks, as shoulder peaks no longer have a measurable derivative peak height. Noise and increased peak tailing also reduce the measurable peak height of the derivative. Other reasons, such as the use of a variable boxcar-averaging window, reject this approach for quantitating overlapped peaks.

One other use of derivatives is to enhance peak resolution of tailed peaks. Here Ashley and Reilley[81] reduced the degree of overlap for tailed peaks by adding linear combinations of the second, fourth, and sixth derivatives of the chromatograms. These sharper peaks were then better resolved, which resulted in increased accuracy when the perpendicular drop method was used to resolve the peak areas. Unfortunately, no explicit function for the correct portions of each even derivative could be derived, and many times too much de-tailing left a derivative-like chromatogram. The effects of noise were not investigated. This method of sharpening peaks has not been investigated further, although the Fourier transform method of sharpening peaks discussed later relies on the same principle.

Grushka and Monacelli[82] used second-derivative extrema ratios to recognize severely overlapped peaks, where the sensitivity of the second derivative made it preferable to the first derivative. Calibration data, in the form of second-derivative ratios, are first required at several known levels of resolution for different peak-height ratios and peak tailing. When this calibration curve is applied to real data of visually indistinguishable, overlapped peaks, calculated resolution was within 20% of the simulated values. Reproducibility was within 5%. Noise effects were not investigated. Although this method requires calibration, it can be used to determine peak purity. If other peaks are shown to be present, the resolution among peaks can be obtained. Reliance on calibration data that may not be representative of sample chromatograms restricts this method as well as others. It is not generally used in commercial systems.

In summary, derivatives are widely used for peak detection and in some cases indicate peak purity. Derivative methods are not generally applicable to resolving and quantitating peaks, however. Noise adversely affects derivative calculations no matter which differential method is used. Noise minimization can be done in several ways and they are discussed below.

D. Curve Fitting

Curve-fitting techniques are perhaps among the most accurate chromatographic curve deconvolution methods that do not require calibration standards. Like every other technique, however, curve fitting relies on some basic assumptions; the number of peaks present in the overlapped group must be known (peak purity), and good estimates of the true peak-shape parameters are required. When these are inaccurate, results are not only inaccurate but invalid. Sometimes even when the peak purity is known and initial estimates are accurate, incorrect solutions may still result.

Curve fitting became feasible in the late 1960s and early 1970s as more computing power became available to chromatographers. Initially, preliminary work was conducted on mainframes in a batch mode. Today, this technique is feasible in an interactive mode on small computers in the laboratory.

Curve fitting is performed by mathematical regression techniques; modeled peaks, shapes, or curves are fitted to the overlapped peaks using nonlinear least squares. The principles were demonstrated visually by Butterfield[83] in which an analog computer of operational amplifiers was designed to generate Gaussian profiles. These profiles as well as the real chromatogram were fed to a multichannel oscilloscope. The analog computer was adjusted until the simulated data matched the overlaid real chromatogram. A "best fit" index is generally used to demonstrate the goodness-of-fit. Here the criterion of "best fit" was the human pattern-recognition process. Peak areas were then obtained by sending the individual waveforms of the analog computer into an integrating operational amplifier. Provisions were made to lag the Gaussian profile to mimic peak tailing. Although not automated nor statistically sound, this work characterizes the process and goal of curve fitting.

Some of the general problems with curve fitting are apparent in the analog pattern-matching technique. These include the age-old problem of finding a general function against which all data is matched, good initial estimates for the parameters in the model, peak end point detection, baseline subtraction, and noise.

Any assumptions about one general function representing all real data lead to problems when the data are not best explained by this model. With curve fitting, additional concerns result from the application of the mathematics. More complex models require more parameters, and each additional parameter leads to an N-fold increase in the number of iterations required. More parameters also make divergence more likely.[84] Additionally, to include separate models for different peaks being deconvoluted requires much more computation. For these reasons only one general functional form is generally used to curve-fit overlapping peaks.

Iterative curve fitting has been performed with the EMG,[59,85] Gaussian,[53,68,86,87] Cauchy,[53] Cauchy/Gaussian combinations,[54] and a combination Gaussian/hyperbola/exponential.[55] Chesler and Cram claimed this latter function is most representative of real data and were not intimidated by the need to fit eight parameters for each peak. Hippe et al.,[87] on the other hand, admitted the EMG is most representative of real data, yet required too many iterations for its three parameters. He chose the two-parameter Gaussian profile. More parameters lead to a more representative model, but unfortunately also make the region of convergence more shallow which can result in slow or indefinite convergence and possible divergence. The same problem of conforming all real data into one general function arises again, only with additional considerations.

Curve fitting only concerns the overlapping peak envelope. Therefore, prior to curve fitting, the peak start and end points must be determined and the baseline must be subtracted out. Any random error or variability here becomes amplified in the final results. Baselines can be assumed to be approximately linear within the peak envelope with little error, although a spline interpolation has been used[86] for nonlinear baselines. Also as noise and peak tailing become prominent, end-point detection becomes more variable.[16-19] Although most peak resolution methods suffer from variable end-point detection, its effect is amplified here.

Noise does not adversely affect the least-squares regression techniques. Noise spikes within the peak envelope present special problems to so-called shortcut methods. These methods have been used to reduce processing time while giving a close approximation to the least-squares procedure.[53,85] The functional model that is to be fit to the data is reduced to its Taylor series expansion, and use is made only of the zero- and first-order terms. The first-order partial derivatives are then calculated analytically before any iterations. Whereas least-squares procedures find the best predicted curve among the positive and negative noise spikes, this shortcut method works with noiseless, nonredundant data. Although shown to closely approximate the least-squares results in limited cases, this seems unnecessary with the computing powers of today.

The last troublesome area encountered is slow convergence, or even worse, divergence.

From the initial parameter estimated and through an iterative process, the parameters are refined, based on the criteria of least-squared error. Most have used the Newton-Raphson local linearization process.[55,85,87] If the initial estimates are poor, the process may converge on false values or never converge. So, in addition to preprocessing the raw data, good initial estimates are needed.

According to Roberts,[86] initial estimates can be only about 25% different from the true values to avoid divergence. His initial estimates were calculated manually. Gladney et al.[85] made initial estimates part of a completely automated curve-fitting process, but did not mention the specific process. Hippe et al.[87] used a Hooke-Jeeves algorithm for getting initial values. Even though different approaches have been taken for parameter estimation, none are completely automatic. This problem remains a stumbling block to curve fitting.

The performance of this technique is impressive when considering the inaccuracies of the presently used perpendicular drop peak-resolution method. Yet it still is not without bias. A comparison of bias among the shortcut method, least-squares curve fitting, and the perpendicular drop on three overlapped peaks was done by Gladney et al.[85]:

Shortcut:	−1% to +12%
Least-squares:	+3% to +12%
Perpendicular drop:	−40% to +30%

Peak-tailing factors were not given. In another study, bias was −5% to +10% when tau/sigma ranged from 1.0 to 1.6 for three overlapped peaks.[59] Equally important, the repeatability was 1 to 2% relative standard deviation (RSD). Although these numbers are not zero, the results reflect values that are better approximates of the true values.

An adaption used by Anderson et al.[84] avoided the problem of representing all data with one general function. A calibration chromatogram containing only one of the overlapped components was needed to define the nonfunctional template representative of all components present. Curve fitting was applied as above. Resolution as little as 1.5 led to accuracies of 99%. However, as resolution deteriorated, the match to the calibration peak profile became critical. A poor match led to divergence, which in some cases was completely undetected; residual analysis showed only randomness. This is not true with modeled data; divergence is detected by a large difference between the initial parameters and the final parameters as well as nonrandom residuals. Undetected divergence leads us to question the reliability of this curve-fitting technique when applied to general data.

Curve fitting has always been the most promising curve resolution technique for moderately overlapped peaks. It is inherently more accurate than all other techniques not requiring calibration standards. The major problem areas include: obtaining better initial estimates of the parameters, false value detection, selecting a general model for all data, precise end point determination, divergence, the lack of a unique solution as well as the age-old problems of noise. Work continues in these areas.

Iterative nonlinear least-squares regression is required to deconvolute overlapped peaks. For curve fitting to be applicable in the general case, the process must run unattended. It is doubtful that any multiple regression technique could be completely automated reliably. Therefore, curve fitting may only appear in future research-grade data systems.

E. Other Peak-Separation Methods

Peak-separation methods briefly mentioned here require either multiple independent channels of data or calibration standards. As such they are not directly amenable to present or near-future, general-purpose laboratory integrators. However, these methods are included to preview what may become available with computer-interfaced, hyphenated techniques such as GC/MS, GC/FTIR, HPLC/photodiode array, and HPLC/ICP.

Goldberg[88] proposed a linear regression technique that required chromatograms of all the

individual compounds within the fused group. Precision chromatographs are a necessity here, as no time or peak shape drifts are allowed. Also, no unknown compounds can be present. With these requirements, the fused group is treated as a linear combination of the individual compounds. Least-squares linear regression then solves for the coefficients corresponding to the weight fraction of each peak. Goldberg used the Gram-Schmidt approximation for computational speed in place of the least-squares procedure, as they result in the same solution. Results were accurate within 99% for known mixtures.

There are many assumptions and inconveniences with this linear combinations approach. The peaks cannot drift in time or shape, so the analysis must be conducted with precision temperature control and/or strict mobile phase composition control. The concentration of the calibration standards must be close to that of the samples to avoid peak shape changes prevalent within wide concentration ranges.[62] No interaction terms are assumed, which does hold for most HPLC detectors. These assumptions lead to errors when shown to be invalid. Others have observed these deficiences and proposed amended versions of this technique in order to overcome some of these problems.

Lundeen and Juvet[62] recognized nonlinearity of response with concentration and included a second-order term. From simulation studies, the inclusion of nonlinear terms led to bias ranging from 0 to 6% for overlapped peak resolution of 0.5 to 0.1, respectively. Bias increased with increasing noise. The authors claimed that this method is theoretically applicable even for completely unresolved peaks that only vary in peak width. Like the original paper, this approach requires calibration standards, no unknown peaks, and a precision chromatographic system. These requirements exclude it from consideration for complex matrix samples with many unknown peaks and few pure calibration standards.

Gourlia and Bordet[89] applied correction factors to the Goldberg method that take into account response differences among compounds. They demonstrated increased accuracies with the use of these correction factors. The suggestion was made to adopt this method to save analysis time; compounds could overlap severely within the chromatogram, yet the mathematics could accurately quantitate the results. Unnecessary time to baseline-resolve these compounds would therefore be saved. Not included in this evaluation, however, was time to elute each calibration standard, time to process the data, and finally the need and expense to operate a precision chromatograph. Chromatographers may not feel comfortable about reducing the chromatographic resolution and letting the separation be performed mathematically.

Another technique applied to fused-peak deconvolution is the transformation of the peak envelope into a series of orthogonal polynomials, both Hermitian and Chebyshev. The entire fused-peak group is reduced into a handful of orthogonal polynomial coefficients, thus reducing data storage requirements. These polynomial terms are not, however, described easily in conventional chromatographic terms, making the peak moments difficult to extract. This approach may be useful in deriving a nonconventional general peak-shape model where addition or subtraction of higher-order orthogonal terms can tailor the series to the chromatogram. This method was shown to be very imprecise when applied to noisy data.

Related to the Goldberg approach is that of Macnaughton et al.,[91] where two or more chromatograms of overlapping peaks at different peak-area ratios are required for calibration. Again, precise control of the chromatography is required as time synchronization and constant peak shape are mandatory. Here, however, the data matrix is transformed into an orthogonal data space by principal components analysis (PCA). From the transformed eigenvectors, pure peak profiles are reconstructed. The accuracy of this method for two peaks can be 99% or more, but uncorrectable time drifts and noise decrease accuracy considerably. This is only one example of PCA being applied to multivariate data.

With multiple channels of data available from hyphenated techniques, monitoring one channel of information is wasteful and may be both imprecise and inaccurate. The intro-

duction of multivariate techniques is necessary to manage and interpret this information simultanously. Factor analysis, similar to PCA, is widely used to determine the underlying peak purity.[92-97] Combinations of factor analysis and library-searching techniques were used not only to determine peak purity, but also to identify the unknown impurity whose spectrum was reconstructed from the library data.[96] The use of multivariate techniques in analytical chemistry has increased dramatically in the last 10 years, as evidenced by the continued growth in curve-resolution applications.[92-95] These techniques will appear in future data systems.

The mathematical techniques briefly mentioned here generally require more chromatographic information before being applied. Single-channel detectors and single injections are insufficient. However, if these techniques are proven to be superior to the present geometric methods and the additional accuracies are demanded, sample analysis protocols and/or instrumentation will adapt. This is especially true of complex samples containing many unknowns.

A general trend within analytical chemistry is the shift from an instrumentation science to an information science. More real-time instrumental data is now provided. A pencil, ruler, and calculator are inadequate to resolve scores of overlapped peaks or to digest data obtained from the hyphenated instruments. Multivariate techniques will appear in future instrumentation. One can only conclude that 5 to 10 years from now, data systems will also make available multivariate techniques for curve resolution and for newer chromatography, such as third-order chromatography.[97]

F. Conclusions

Curve-resolution techniques vary widely in simplicity and computation requirements. The need for more accurate methods beyond the simpler geometric method has always been present. However, every proposed technique must make assumptions that cannot always be true, making them less reliable than the geometric method. It is not likely that some new, simple technique will be discovered beyond what currently exists. More likely, improvements will come from continued enhancement and refinement of the present methods. As computing power at the bench increases, more sophisticated techniques will be used. For the present, though, the perpendicular drop method remains the only practical method of choice.

III. NOISE AND FILTERING

A. Sources and Characteristics of Noise

Every analytic measurement involves distinguishing a signal from background noise. The level of the signal compared to the noise is described as the signal-to-noise (S/N) ratio and defines the quality of the analytic measurement. Many characteristic noise sources contribute to the overall observed noise. Their relative contribution depends on the analytic method, the detectors used, and the operating environment. Different analog and digital means of reducing noise are individually optimized for a particular frequency bandpass. These filters are used both physically within the detectors and electronically within data systems. Applying the wrong technique or maximizing the noise rejection leads to overfiltering the analytic signal and distorts the peaks.[98-111] To introduce minimal distortion and error while extracting maximal information, knowledge of noise characteristics and optimized filtering is required.

Noise cannot be easily characterized in the time domain as its periodic waveform is unpredictable. Transformation into the frequency domain is required. A power-vs.-frequency spectrum (power spectrum) points out the frequency dependence typifying the different noise sources. The noise power spectrum can then be compared to the power spectrum of the signal, revealing the strategy used to reduce particular sources of noise.

Uncorrelated white noise is composed of all frequencies as seen by its uniform distribution

throughout the power spectrum. The origin of white noise is Brownian or random motion within charge carriers and appears random in phase and amplitude when viewed in the time domain. White noise is the most predominant source of noise in many analytic methods. Although a source of white noise such as a single resistor has a uniform distribution, several resistors are typically encountered within one detector. The aggregate of these uniform distributions sum to a Gaussian distribution by the Central Limit theorem. This becomes important when simulation studies are conducted that require correct models of white noise.

Interference noise appears as sharp spikes at the primary frequency and also at its harmonic frequencies. The most common interference source is from the power line, resulting in spikes at 60 Hz and at multiples of 60. Other sources include high-voltage transformers, radio transmitters, and the high-frequency backplane of a computer. Because of its narrow frequency bandwidth, interference noise can be easily filtered out.

Impulse noise appears as a sharp disturbance in the time domain, is usually caused by a sudden drop in the line voltage, and is seen frequently when a motor is started. In the frequency domain impulse noise has a broad spectrum, very similar to a white noise spectrum. Because of the frequency independence and its transitory existence, impulse noise is the most difficult to eliminate.

Another source of noise commonly seen in chromatography is 1/f noise or drift. Drift can fluctuate from low to high periodicity. Slow drift is easily corrected within chromatographic data systems by baseline subtraction techniques, where a linear approximation is used. When drift varies rapidly enough to make it indistinguishable from wide chromatographic peaks, noise filtration techniques are inappropriate and must not be applied; this situation generally indicates an instrumental problem. Data systems do not try to filter out drift, but instead use baseline subtraction techniques.

Although other characteristic forms of noise exist, they are not treated here. Laeven et al.[112] developed software to simulate first-order, Gaussian, and dampened-cosine, first-order noise as well as the above-mentioned white, 1/f, and impulse noise. All these noise types were used in evaluating one peak-detection algorithm. These three additional forms of noise are generally the result of some electronic preprocessing of the signal by operational amplifiers circuits; the bandpass characteristics of the circuit define the observed preprocessing noise observed. To a large degree this preprocessing is not used in chromatographic detectors and need not be considered important.

B. Noise Reduction

Reduction of white noise can be performed in many ways, depending on the frequency of the signal. In Fourier transform infrared (FTIR), several spectra are averaged. In atomic absorption, modulation is used. Autocorrelation techniques are used for periodic waveform methods, such as AC polarography. Boxcar integration is another commonly used method. In averaging techniques, the random positive- and negative-noise spikes cancel each other out while the signal adds constructively.

Data systems use several forms of averaging to reduce white noise. Within the hardware, many systems use an integrating type analog-to-digital converter (A/DC), such as a voltage-to-frequency or a dual-slope type A/DC. This A/DC performs the same function as a boxcar integrator; the signal is summed over a finite time period, allowing the positive- and negative-noise spikes to cancel while the signal remains at a constant DC level. This is the first step toward noise reduction.

White noise reduction can be further reduced by boxcar averaging data points. The window or boxcar width must be variable so narrow peaks are not overfiltered.[102,103] When done in software, this is called data bunching. Some commercial systems that use discrete A/DCs at a higher sampling rate (such as the successive-approximation type) have a nonadjustable low-level averaging routine. The bunching window is adjusted by the peak-width parameter.

By using a variable window, not only are noise levels reduced, but also the effective slope sensitivity is increased. This sensitivity increase occurs because slope is based only on the net change within this window; as the number of data points increases, the net change becomes larger.

In data systems that perform reintegration, it is bunched-data points that are reintegrated and not the original raw data.[112] This approach will not give correct peak quantitation if the peaks are overfiltered and distorted by an incorrect peak-width value; sharp peaks are rounded out by the averaging process. The peak-width value must be based on the narrowest peak in the chromatogram, and the resulting integration must be checked by the chromatographer.

Interference noise generally can be eliminated by notch filters and by sampling at exact intervals of the interfering frequency. For example, sampling at 1/60 s will cancel out the line power interference noise; noise reduction of 1000:1 has been reported.[24] Interference noise generally is not as much of a concern as white noise because its high-frequency band is far from the low-frequency band of chromatographic peaks.

Finally, digital filter techniques act as low-pass filters and eliminate high-frequency noise. Two common examples of symmetrical, linear, nonrecursive digital filters are the moving average (MA) and the polynomial least-squares (PLS) filters. Both methods attenuate random Gaussian noise and attenuate high-frequency spikes. Many data systems use both techniques.

Much has been written on the Savitzky-Golay[79] PLS filter[102-108,111] which remains the standard against which others are compared. Comparison of the MA and PLS was performed by Cram et al.[104] The PLS filter introduces less peak distortion than the MA when the same convolution window is applied, yet it is not as effective in noise reduction. This is especially true when multiple smooths of smaller convolute intervals are used. Small convolute MA filtering is adequate as a preliminary noise filter, whereas a larger convolute PLS differential filter is desirable as a final step. This is how several data systems filter and differentiate the signal.[104] As peaks increasingly tail, more effective noise filtering is performed by smaller convolute intervals of a higher order. Proposed hardware PLS filters[115] would allow neither the window nor the polynomial coefficients to vary. This is partially achieved by the software; the variable data-bunching rate effectively varies the convolution window. Few commercial algorithms switch to a higher-order polynomial.

Few data systems use PLS filters for strong filtering. Yet to yield a maximal S/N enhancement factor of about four, distortion approaches 17% in peak-height and -width measurements, whereas peak-area distortion is about 1%.[201] With strong filtering of differential spectroscopic signals, the oscillating roll-off apparently causes distorted double peaks to appear where single peaks exist.[105] Various amended PLS filters have appeared[105-108] that either dampen out the roll-off, offer a recursive algorithm, or change the form of the window. Enke and Nieman[102] pointed out the need for proper sampling density to avoid distortion. Thompson and Dessey[103] pointed out the need to avoid quantization errors. These filtering concerns are not generally a problem with commercial chromatographic data systems where algorithms avoid strong filtering. However, filtering noise without distortion of the signal can be a problem for those who choose to develop their own integrator.

C. Simulation

Simulation of noise representative of real chromatographic signals is useful when evaluating the performance of electronic integrators. Some advocate simulating several types of noise beyond what is normally encountered.[111] However, the same author (Miscoski), when evaluating the performance of one peak detection and quantitation algorithm, found virtually identical results among six types of simulated noise. Most evaluations have only been concerned with white noise. Even in this case, however, only a few have simulated the noise of multiple sources (Gaussian amplitude distribution)[101,107,111,112] instead of single-source white noise (uniform amplitude distribution). Although this would appear to be a

minor point, and Enke and Nieman[102] demonstrated that PLS filtering is much more effective for uniform noise than normally distributed noise. Thus, in evaluating the peak-to-peak noise effects of PLS filtering devices such as electronic integrators, the use of uniform noise would demonstrate better noise immunity than truly exists.

D. Peak Sharpening

Noise filters generally distort chromatographic peaks into shorter, wider peaks of reduced resolution. Filtering sometimes causes a sharpening of the peaks, enhancing resolution. Bromba and Zieglier[107] proposed the use of a two-parameter, recursive filter that sharpened peaks in some cases. The enhancement was not large and in some cases nonexistent, yet the filter did not distort peaks either. Others have proposed the application of the fast Fourier transform (FFT) to enhance peak separation.[109,110] However, in both FFT cases, assumptions were made about peak shape. The first case assumed a common Gaussian peak shape of equal variance to be effective.[109] When the assumptions failed, so did the resolution enhancement. The other[110] assumed a constant peak shape that limited the method. Computation times for FFT[114] do not make these FFT enhancements practical, considering the limited resolution enhancement. Caution must always be heeded as any resolution enhancement can cause the appearance of spurious features. However, as work continues, refinements such as the work of Bromba and Zieglier may permit peak sharpening instead of peak distortion, while reducing noise.

E. Additional Concerns

There are several other factors to consider when operating with digital logic, i.e., computers. These factors include sampling rate, resolution of the hardware, order of arithmetic operations, and significant figures.[100,101,115,116] For those who develop their own computer-based integrator, all of these factors are important and must be given careful consideration. In commercially available systems, only the relative sampling rate can be changed by the chromatographer. Chromatographers must assume that the other factors have been optimized by the manufacturer. Briefly mentioned here is the relative sampling rate.

As mentioned above, the data aquisition hardware of most commercial systems consists of an integrating-type A/DC, typically of 16-bit resolution. With these A/DCs, the signal of the detector is effectively integrated, then transferred to the memory of the computer as a discrete data point; the rate of "sampling" is the rate of polling. These data points are bunched and then digitally filtered. The rate of polling is on the order of 60 to 100 points per second. It has been shown that the sampling rate that results in minimal error is 50 points per sigma (Gaussian peak width).[100] Although symmetrical peaks can be sampled fewer times, tailed peaks require higher rates. High sample density was also shown to be important in avoiding distortion with PLS filtering.[102] Oversampling, on the other hand, leads to increased spurious noise. This effect will be most dramatic for narrow peaks and will set a lower limit for fast capillary peaks. Based on the criteria of 50 points per sigma, inaccuracies will be incurred for peaks narrower than 4-s basewidth. The ability of several commercial integrators to quantitate capillary peaks accurately was evaluated and is discussed below.

These concerns are also relevant when simulating chromatography; the sampling rate of the simulator should be greater than the rate of polling of the data system to avoid frequency aliasing. If the sampling rate is low, the variance observed is the sum of that due to the data system under test and that due to the simulation process. The contribution due to the simulation process depends on the type of A/DC within the integrator and its sampling rate; this contribution will not be constant and cannot be simply subtracted. Therefore, effective simulation studies require sampling rates greater than that of integrators.

With either the wide dynamic range integrating A/DC or an auto-scaling discrete A/DC

of 16-bit resolution, the minimum hardware resolution or the least significant bit (LSB) represents 0.0015% of full scale. When using 12-bit resolution hardware available with some commercial systems, one LSB is 0.024%. Cost is a primary consideration in selecting the lower-resolution hardware. However, several present (computer) era peak integrators[21,22,26,27,29] have been developed with 8-bit A/DC converter hardware, presumably due to the additional complexity of programming a 2-byte transfer with these 8-bit microprocessors. With full-scale peaks (relative to the range of A/DC) the minimum resolution of a peak is 0.4%, limited by this hardware. With half-scale peaks this becomes 0.8%, quarter-scale peaks are 1.6%, etc. The quantization error of trace peaks could exceed 40%. This level of inaccuracy is not acceptable with any quantitative method of analysis and should be used only for screening models.

IV. EVOLUTION OF PEAK-SENSING DEVICES

The evolution of peak-sensing devices required the advancement and improvment of solid-state logic devices for faster and more complex algorithms. Earlier commercial integrators were shown to lack adequate peak-detection logic for typical chromatography,[3,117-122] demonstrating the need for improved peak quantitation. As with any field, refinements were made incrementally and may only appear subtle or nonexistent to the average user. However, a major advancement occurred when computers were applied to this problem. Earlier applications utilizing computers were handicapped by relatively slow and expensive memory, slow processor speeds, and expensive data acquisition hardware. Much of the postrun calculations were done in a batch job. Large-scale integration made computers more compact, powerful, and affordable. Many present-day PC-based data systems outperform the earlier mainframes. Ironically, algorithms used by these newer, more powerful systems use the same (refined) peak-detection logic as the earliest vacuum-tube integrators.

A chronological review of peak-sensing devices is presented in this section; vacuum-tube devices gave way to solid-state operational amplifiers. The vacuum-tube peak integrator patented in 1958 by Bohm[125] led to the first commercial electronic integrator in 1961.[2] Transistor logic circuits with memory storage displaced these devices, and the first computing electronic integrator was produced in 1971.[7] Time-sharing mainframes led to benchtop PCs. Figure 9 demonstrates this chronological development.

To demonstrate the improvements made with each new generation of devices, several difficult chromatographic situations are shown in Figure 10. The ability to deal effectively with each of these conditions represent an incremental improvement in peak-detection and baseline-subtraction logic. However, these three situations are not the only conditions known to cause difficulty with chromatographic integrators. Evaluation of chromatographic data systems is discussed in the next section. For the purpose of demonstrating advancements over time, handling these situations well represent milestones in the development of chromatographic integrators.

A. Manual Methods

The most basic manual methods for peak integration are the graphic geometric methods (Figure 4). Although there is nothing automated about measuring peak area by triangulation, planimetry, cutting and weighing, nor peak height multiplied by peak width, they are presented for completeness. Prior to 1961, these were the only available methods to determine peak areas.

For the simplest case of peak height times peak width, Ball et al.[14,71] investigated the three relative measurement errors of baseline, height, and width. As expected, the precision obtained depended on the shape of the peak; sharp peaks, small peaks, and broad peaks represented the most difficult cases. Sharp, small, and broad peaks will always represent

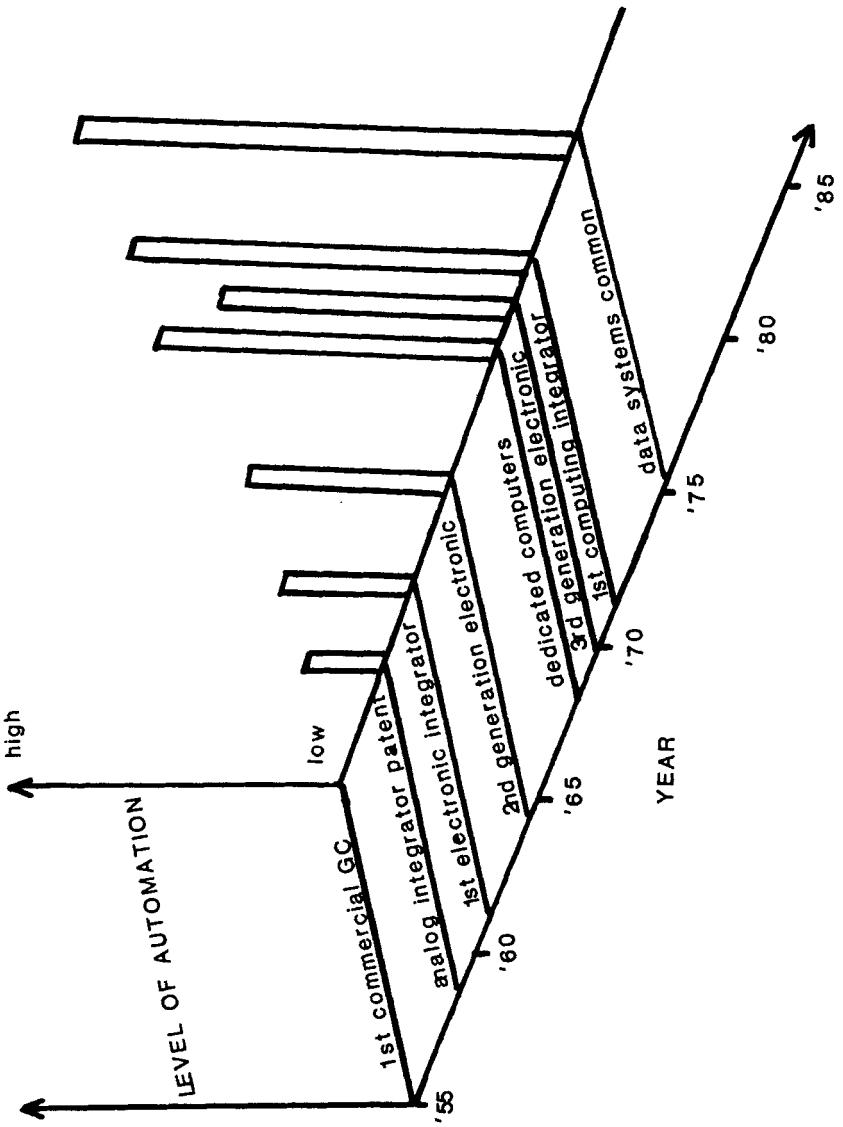
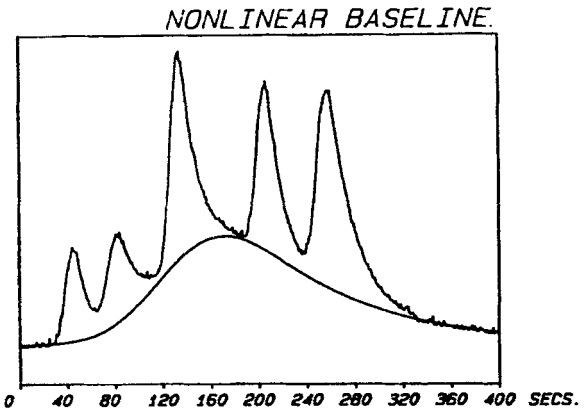
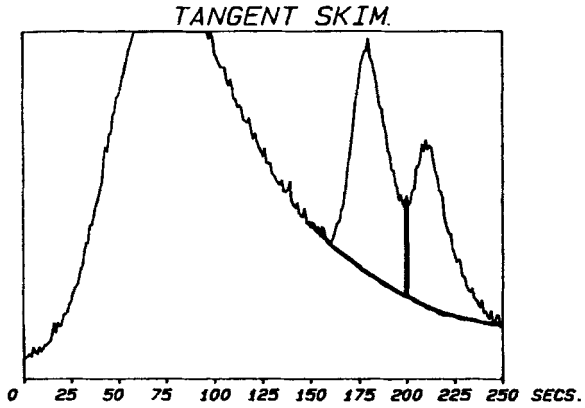


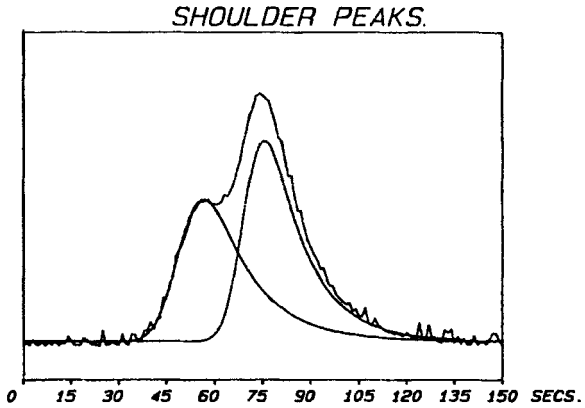
FIGURE 9. Chronological development of peak integrating devices.



A



B



C

FIGURE 10. Three difficult chromatographic situations. (A) Nonlinear baselines that cannot be accurately described by linear, nonzero approximations. (B) True, nonlinear tangent skimming. (C) Shoulder peaks.

problems, no matter what method is used. Relative errors varied parabolically from 1% to more than 10% for peak shape that went from sharp to wide. Peak tailing was neither specified nor controlled, and noise levels were minimal. In a study by Janak,[123] similar results were found. When simulated EMG peaks were measured by this method,[124] the relative error in measured-peak areas increased linearly with increasing tau/sigma values. For example, at tau/sigma of 1.0, the relative error was 8%, and at tau/sigma of 3.0, the relative error was 19%. The Condal-Bosch variation of this method (peak height times the average of the width at 15 and 85% of height) proved almost insensitive to peak tailing in this study and would appear to be a more accurate method.

Relative peak-area errors for triangulation were similar to those of the height/width method in Ball et al.'s study,[71] in Janak's study,[123] and in the study by Delaney.[124] Only in ideal cases were the relative errors less than 2%, whereas with sharp peaks the relative error was about 20%. This was attributed to the measurement process itself by Ball et al.,[71] yet this work did not take peak asymmetry into account. Planimetry was barely an improvement.[71,123] Similar results were obtained for cutting and weighing.[71,123]

According to Janak in 1959, these manual methods were assumed to contribute only 1 to 3% relative error. When shown otherwise, he concluded that the (then) current methods were not optimal. Because of this fact and the need for a more automated method, many devices were constructed to perform peak integration.

B. Analog Methods

Although the first electronic integrator appeared in 1961,[2] many mechanical, electromechanical, electrochemical, and analog integrators were devised and tested. To include all of them would be impossible; however, a few representative ones are mentioned.

Janak, in his study cited above,[123] also evaluated three "automatic" peak detectors. The first was a mechanical integrator consisting of a rotating wheel and a timed cam wheel, with no peak-detection nor baseline-correction logic. Relative errors in determining peak areas varied from 36 to 2% as peaks widened. The second device was an electromechanical integrator consisting of a DC electrometer that changed with the analog output of the detector. Relative errors in peak areas were from 10 to 1.3%. Finally, an analog integrator gave the best results of 2.2 to 0.9%. This analog integrator was constructed similarly to that patented by Bohm[125] in 1958. It consisted of a differentiating circuit for peak detection and a voltage-to-frequency converter. When the first derivative of the signal became nonzero, the analog voltage was converted to a frequency-pulse train that was summed. The counter continued to sum the frequency pulses as the derivative went negative and stopped when the derivative returned to zero. As mentioned above, this derivative peak-detection logic used originally in vacuum-tube systems is still used today.

Ball-and-disk integrators appeared in 1959.[120] These devices provide integration of the signal, but have no peak detection logic nor any baseline-correction mechanism. In a limited study of precision using rectangular waveforms, the RSD of ball-and-disk integrators was 0.5%.[119] With a four-component GC mixture, the RSD for determining peak areas was over 2%. In practice, precision of the ball-and-disk integrator was governed by the operator. Johnson[118] pointed out that these devices are limited by the time constant and dynamic range of the strip chart recorder.

Strickler and Gallaway[126] developed a peak-detection mechanism for this device. The rotational motion of the disk drove photosensitive paper past an intense light source. As the peak passed through the detector, the disk was driven at maximum speed, allowing little light to reach and expose the paper. When the peak came to an end, the light heavily exposed the paper, leaving a sharp line. The distance between the sharp lines was proportional to the peak area. The resulting pattern was termed a "strip integram". No evaluation was made on its performance. This peak-detection method was only one of several attempts to automate the ball-and-disk integrator.

Sawyer and Barr[119] also compared a DC fixed-field integrator and an electrochemical diode integrator. The electrochemical integrator was found to be the most imprecise device in the study with a RSD for peak areas of 4.5% using rectangular waveforms. On the other hand, the DC fixed-field integrator was found to be the most precise in the study. Neither device employs any peak or baseline logic. Other analog integrating devices include an inertial magnetic switch mounted on a recorder[122] and a friction clutch on the pen motor;[122] both of these devices are very poor in accuracy and precision.

C. Digital Integration with Analog Logic

In the late 1950s and early 1960s no device was capable of high precision and high accuracy except for the analog integrator. None had automatic peak detection and baseline subtraction, except the analog integrator. Solid state voltage-to-frequency (V/F) converters appeared during this time period and became widely used for conversion and peak integration. In operation, V/F converters continuously transform the analog voltage to a square-wave pulse train whose frequency is proportional to the voltagle level. A digital counter, triggered continuously by the leading edge of the square wave, is started and stopped by the peak-detection logic or by hand. The accumulated count is read by a Nixie tube. For more than 20 years, tremendous effort has gone into developing and refining the peak-detection logic and the baseline-correction logic of the V/F converter. In fact, the V/F converter conversion and counting scheme is still used in modern-day integrators.

Although digital counters were inexpensive in the late 1960s, integrated digital logic circuits were not. Hence, all logic was performed with analog operational amplifiers. In principle and in hardware, there were two distinct methods for peak detection[117] (Figure 11). The first method would differentiate the signal and use any positive deflection to start the V/F converter. The second method would just wait until the signal exceeded a preset threshold to start the V/F converter. The first method was shown to trigger falsely by noise, whereas the second could not function with nonhorizontal baselines (see levels 1 and 2 in Figure 11).[117] Additionally, neither unit could subtract out nonzero or nonlinear baseline contributions, perform tangent skimming, or recognize shoulder peaks. These first-generation electronic integrators were only appropriate for baseline-resolved, symmetrical peaks with little noise and little drift;[121] yet these integrators were used widely, being the most precise, accurate, and automatic devices of the time.

Karohl[122] pointed out that the first derivative peak-detection method was superior to the level detection scheme, but also needed nonzero, noise-rejection threshold logic. This is exactly what was done in the next generation of commercial electronic integrators[2] to reduce noise susceptibility and to allow shoulder peak detection. Horizontal baseline contributions were also nulled out. Others pointed out that this mode of noise rejection lowered the slope sensitivity, resulting in peak integration that started too late. Because the analog circuitry could not go back and determine the correct peak starting time, this approach led to peak-area inaccuracies of 0.1 to 10%.[121] Alternatively, some manufacturers decided to reject peaks after integration by a minimum area requirement rather than reduce the slope sensitivity. Fused peaks in both cases were resolved by the perpendicular drop method, and neither unit was capable of compensating for nonhorizontal baselines or capable of performing tangent skimming. These features became available only as memory became available; they appeared later in computing electronic integrators and data systems.

Here was the first of many cases where manufacturers agreed in principle on the basic approach to a problem, yet differed in the slight details that define algorithms. As time progressed, these refinements became less publicized. This can be demonstrated by the large number of publications describing new products, yet excluding any information on actual performance, evaluation studies, or comparisions.[127-138] To demonstrate how complex algorithms can be, one can review the peak-detection scheme proposed by Woerlee and Mol.[147]

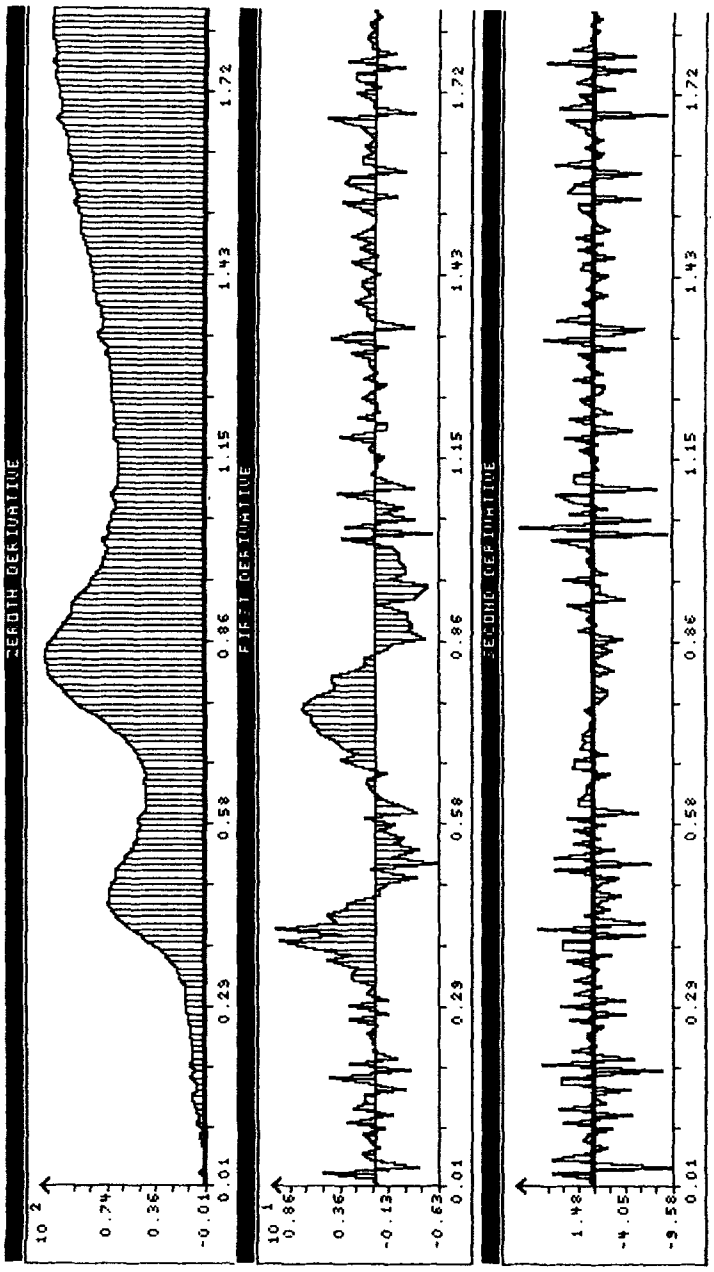


FIGURE 11. Two peak-detection schemes. (A) Differential method is unaffected by baseline drift, but has difficulty discerning noise from peaks. (B) Level detection can reject noise on a horizontal baseline, but neither level 1 nor level 2 allows peaks to be discerned from baseline drift.

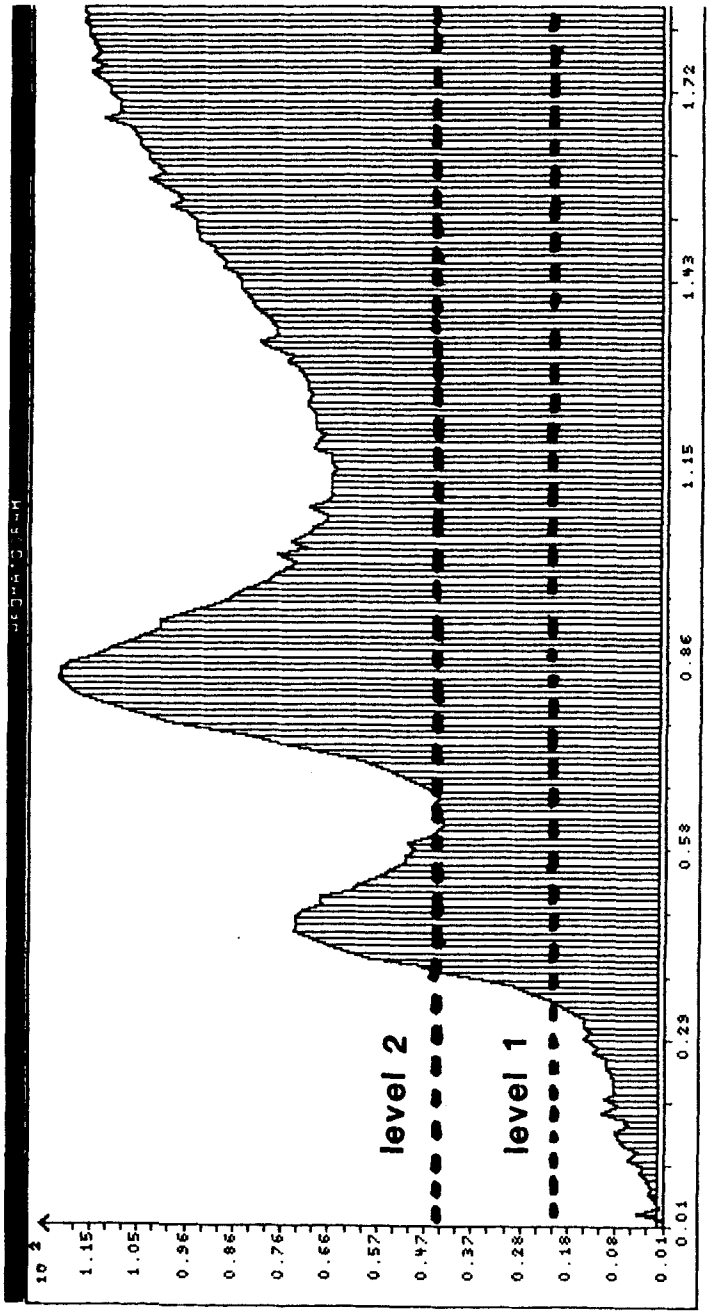


FIGURE 11B

All of these subtle differences lead to algorithms optimized for particular types of chromatography; one integrator cannot be optimized for all chromatographic encounters. A performance evaluation is needed to assess how well an integrator can handle a given chromatogram. Evaluation techniques are described below.

D. Digital Integration with Digital Logic

Although computers were used to analyze chromatographic raw data as early as 1963,[118] the change from analog logic circuitry to digital computer logic was slow. This was primarily due to the cost of computer time, data acquisition hardware, memory, data storage, and limited knowledge of computer operations. Reviewed are hardware techniques used to get the chromatogram into the computer, the types of computers utilized, and the level of software automation.

Before on-line peak detection and quantitation, computers served in an ancillary role. All of the earlier analog integrators had peak-detection logic and provided a printout of retention times and peak areas. None of the early analog devices had sufficient computational power for calculations such as internal and external calibrations, response factor scaling, and normalization. Chromatographers would use electronic integrators to gather peak areas, then batch-process postrun reports and calculations.[1,2] Some went to great lengths to obtain these calculations; they teletyped the results to a time-sharing computer and received the results in the mail.[139] Even modern-day integrators without BASIC could not provide enough computing power for postrun calculations and statistics.[140]

Chromatographers took different approaches in getting raw data into the computer when data acquisition hardware was not available or real-time operation was inappropriate, i.e., time-sharing mainframes. Hegedus and Petersen[28] used a digital voltmeter to convert the analog signal to digital data and then transferred the data to paper tape. This paper tape was read by the paper-tape reader of a mainframe, and peak quantitation was done in a batch job. Burke and Thurman[24] had a digital voltmeter tied directly to a parallel digital port so that the intermediate step was unnecessary. Johnson[118] had the digital counter of a V/F converter continuously feed data to the computer, and the software performed peak detection and quantitation in real-time. This work was done in 1963 and was a prototype of the procedure used today.

In the late 1960s, computer-automated chromatography was classified into five distinct categories.[1] The *off-line* mode was for the above-mentioned postrun calculations. The *hybrid* system combined a digital computer to perform postrun calculations for data from several analog peak integrators. The *time-sharing* system was capable of analyzing several channels of input, switching between channels under control of a time-sharing executive. The next step was a *multichannel dedicated* system that performed primarily only peak detection and quantitation; due to core memory restrictions, postrun calculations were minimal. Finally, the *dedicated* system had individual A/DC channels and performed both foreground peak detection and background calculations. Each successive level in this progression cost more per channel, and the precision of peak-area quantitation increased at each level. At that time, there were severe compromises of precision for price.

Several mainframes were adapted to chromatographic processing,[141-145] as well as several dedicated minicomputers.[4,6,30,146,147] These approaches varied both in hardware and software complexity. Sampling rates varied from 60 Hz,[24] 70 to 10 Hz, or less.[28,30,147] A/DCs varied from the dual-slope[149] to the successive-approximations type.[70,141,147] Resolution of the hardware varied from 16 bit,[141] 15 bit,[70] 12 bit,[30,149] to 11 bit.[147] Software was admittedly "rudimentary",[24] "primitive",[28] or "not flawless".[30] It was reported that "subtle bugs were uncovered",[4] and the perpendicular drop sometimes gave erroneous results.[142] Additionally, the peak-detection algorithms sometimes required ten,[14] nine,[6] seven,[4] five,[142] or three parameters[146] to correctly quantitate peaks. From personal experience,[76] the more

parameters required to adjust, the more difficult it was to find optimal conditions, leading to poor reproducibility. Thus, the application of computers to peak detection and quantitation was not error-free nor uniformly applied.

In 1972 there were two computing integrators commercially available,[7,148] whereas many data systems soon became available.[3,5,149] By 1976 it was estimated that there were over 150,000 GCs and LCs in use,[150] and many models of integrators/data systems had been introduced.[151] The number of computer automation systems developed in-house diminished as their development could no longer be justified financially. As digital logic became cheaper, commercial microprocessor-based systems became readily available to chromatographers. Extrapolating this explosive growth in economically feasible computer hardware, in 1972 Anderson[148] predicted a single "detect-a-peak" chip. Software has not experienced the same growth, and this prediction has not materialized.

V. EVALUATION TECHNIQUES AND RESULTS

Analytical chemists in general must know if their instrumentation is working properly before analyzing a sample. Similarly, chromatographers must know if the individual components of their chromatograph are functioning correctly. Rigorous quality assurance (QA) programs have been set up to demonstrate proper operation of instrumentation. Neglected under this QA inspection has been electronic integrator/data systems, presumably under the assumption that they have been adequately validated by the manufacturer. Few rigorous means to validate the performance of integrator/data systems within the laboratory have been advanced. Thus, the most critical component of quantitative chromatography remains unchecked.

Today, data systems fall under the classification of computers which directly control product quality,[152,153] and as such must comply with Computer Good Manufacturing Practices.[153] Many standards, guidelines, and books[154-158] have been written on QA for computer software and hardware. Regulated manufacturers can no longer accept black-box devices without assuring adequate performance.[152] Audits of QA programs are beginning to include chromatography data systems.[153] Methods for validating integrators used over the years are presented below.

As with the first analog integrators, manufacturers may agree on the general approach to peak detection, yet vary significantly in the details of implementation. Also demonstrated were wide differences in hardware and software sophistication as well as the attention given to details when computers were first applied to peak quantitation. Many bugs were noted then. Anderson declared in 1972 that development of algorithms had "leveled out", and that most commercial systems are approximately equal.[148] If equal in 1972, then 15 years of refinement should have resulted in flawless operation today. What do chromatographers think?

In 1986 a survey of chromatography data-system users was taken for problems with data systems.[159] Thirty-two percent of those polled cited ease of use or software problems, whereas over 22% of those polled cited erroneous results as their major problem. Erroneous results centered mostly around incorrect baseline selection. Although some of these problems may be due to a misunderstanding of correct operation or insufficient documentation for the chromatographer, it would be difficult to dismiss all the complaints. The need to develop a QA program for integrator/data systems is not just attributable to regulatory requirements, but is also required to demonstrate correct operation.

Over the years advice has been given on how to evaluate data systems before their purchase as well in daily operation. Gill[1] in 1969 recommended chemists evaluate the details of the peak-processing logic, fused-peak algorithms, and calculations from the manufacturer before making a purchasing decision. The best method, according to Gill, was to visit the manu-

facturer and operate the system with one's own gas chromatograms first. Similar advice is echoed in a survey of chromatography data-system users in 1985.[160] After purchasing, Dolan[159] advised users to check the data system for reproducibility and accuracy before injecting samples. These are reasonable suggestions, yet sometimes cannot be followed.

The intimate details of proprietary algorithms are not usually available. Even if they were, the relationship of an algorithm on paper to the results of real samples is not obvious. Manufacturers are, however, usually willing to lend chromatographers their data systems for purchase evaluation. Chromatographers can set up their equipment and test the integrator with their chromatograms. Similarly, for daily checks one can make several injections to calculate reproducibility. Accuracy can be evaluated only if completely resolved peaks are used in the method. This procedure requires having solvents ready, columns equilibrated, detectors warmed up, and standard solutions all prepared. An easier method can be used and meets the requirements for software QA standards. This general method employs simulation, or a black-box exerciser, as discussed below.

A. Waveform Generators

The first method of evaluating peak integrators requires a waveform generator. Triangular, square,[117] and, in one case, a Gaussian waveform[147] have all been used. That the triangular- and square-wave generators were easy to fabricate was a prime consideration in their use. However, chromatography is not well represented by either waveform, and any evaluation with them is not indicative of performance with real data. Only with the Gaussian generator does the simulator begin to approximate real conditions.

Recent work[161] demonstrated that Gaussian peaks are good only to test the precision and accuracy of isolated peak detection and quantitation. When evaluating fused-peak resolution, the results of Gaussian profiles did not correspond to tailed EMG profiles. From unpublished results, predictable results were obtained when resolution among unresolved Gaussian peaks varied from 2.0 to 1.2. This was not true of fused, tailed EMG peaks at the same levels of overlap; flaws in baseline construction were observed with some commercial integrators. Based on this work, it appears that a Gaussian waveform representing ideal chromatography cannot be used to evaluate performance in the more difficult situations of Figure 10.

It is possible to construct a black-box EMG (or any other function) waveform generator. However, such a simulator would require flexibility and precision beyond what are presently available. Given that chromatography in one laboratory is not representative of another, this device must be capable of generating an infinite variety of chromatographic situations. This would require an interactive device creating chromatograms in real-time. Although feasible, it is not likely that such a device could compete with a laboratory PC using chromatogram-generating software. Thus, a black-box simulator with the necessary flexibility required to evaluate integrators effectively is feasible, but may not be practical.

B. Real Chromatography

The most obvious and predominant approach to evaluate peak integrators is to use real chromatographic signals, whether in real-time or in playback; after recording the real data in some retrievable form, replay it on demand. Widespread use is also made of accurately prepared standards of known concentrations; inject them into the chromatograph, record the resultant peak areas, then construct a calibration curve. Each approach provides valuable information, yet each is subject to constraints that can obscure the properties sought or limit the scope of the evaluation process.

Recording the chromatrogram on magnetic tape can be done easily by using a V/F converter.[25,118] This stores the raw chromatogram as frequency data that can be retrieved later. By playing back at multiples of the recording speed, an increase in effective slope sensitivity can be realized, in one case by a factor of 16.[118] Paper tape has also been used for this

purpose.[139] Although these are inexpensive solutions, there are problems of storage and playback, however.

Magnetic tape has limited dynamic response and limited noise rejection, and requires high precision in both the recording and playback drives.[162] Paper tape is further restrictive in these properties. The minimal level of variance that can be observed while evaluating data systems is constrained by the largest source of error. Using magnetically recorded chromatograms, the largest source of error will not be the data system being tested, but the manner in which the chromatogram was produced. For this reason, this and other ways of generating a chromatographic response are also rejected.

A rather unique way of generating peak profiles was shown by Johnson[162] in 1968. From a real chromatogram, peaks were cut out and placed on an opaque background. Mounted on a rotating drum, this spinning reversed image passed by an array of photo cells whose output dropped as the opaque image passed by. This current was then inverted and converted to a voltage. In this manner, the original chromatogram was reproduced. To demonstrate flexibility, peaks could be added, subtracted, or modified; peak widths could vary as the speed was varied. However, the reproducibility in generating the peaks was shown to be 2%, which is about the level of reproducibility for experimental chromatography. Since integrator variability may be less than these levels, this simulation approach is not given further consideration. This method does, however, point out the advantages of having stored test chromatograms that can be easily modified.

The final means discussed here to obtain test chromatograms for evaluation studies is also the most widely used today. Before real data can be processed by a microprocessor-based system, it must be converted to binary data. This binary data can also be stored and archived. Later this binary data can be fed to a digital-to-analog converter (D/AC), recreating signals that are most authentic. For one manufacturer, this was a useful field-testing device.[163] There are some disadvantages to this approach, however. If a systematic evaluation process is sought, it is difficult to control the chromatography well enough to provide systematic data. For example, to incrementally increase peak tailing for only selected peaks of a particular chromatogram would require a variable dead-volume chamber be applied only to those selected peaks, or, for a series of peak-area ratios, separate solutions must be prepared. Additionally, true peak areas are unknown, not allowing accuracy to be investigated. It is also difficult to corroborate this work with others.

Recreating chromatograms is the predominant evaluation method used by some manufacturers today because it demonstrates performance with real data. Even in this case, it would be advantageous to better characterize system performance with meaningful parameters, such as peak moments, noise levels, tailing factors, skew, etc. Although these relationships can be calculated from the data at hand, a systematic means of altering these parameters is not available. Only precision, not accuracy, can be assessed. Despite these desired features, recreated chromatograms do evaluate performance with data that are sometimes unpredictable. Therefore, playback of stored real chromatography combined with simulation studies is the most informative and efficient means in evaluating performance.

C. Weight Percent Concentration vs. Quantitation

Before hardware simulation became popular, evaluations relied on testing the performance of data systems with chromatograms of precisely known standard solutions.[3,6,24,117,119,141,142,144] In some cases, linearity of peak areas with dilutions were sought,[24,28,142] whereas in others, comparison between sample weight and peak area was the main objective.[3,4,117,119,141,144] In all but a few cases, baseline-resolved peaks with little tailing were used. In precision studies, RSDs varied from 0.2[4] to less than 2%.[3,117,119,141,144] In one test to determine the accuracy for quantitating overlapped peaks, the GC temperature was varied for the same known standard solutions, causing various levels of peak overlap.[144]

Using experimental data for precision studies has some drawbacks. The RSDs calculated reflect the sum of the variances due to all the components. This includes the variance of the solvent or gas-delivering system, the injection, the detector, and the data system itself. In most cases, the variance of the chromatograph is much larger than that of the data system; in simulation studies using 16-bit hardware,[161,163] the simulation process itself introduced a variance of about 0.002% or two orders of magnitude lower than that of an experimental approach. The precision observed in one study[161] for several different manufacturers was also two orders of magnitude lower than the values cited above. If the results of the simulation study reflect the true variance of peak detection and quantitation for commercial systems, then the variances from experimental studies reflect variance due primarily to the chromatography itself.

This can be further demonstrated by examining a collaborative study conducted by the ASTM subcommittee of chromatographic data on the precision of chromatography data.[16] Here two standard separation mixtures were sent to 78 laboratories with recommended methods of analysis — very clean, well-resolved separations. Peak areas were calculated by whatever integrators were available to each laboratory. For the first separation of four peaks, RSDs varied from 6 to 11%, whereas for the second separation of six peaks, RSDs varied from 9 to 16%. It would be foolish to assume all this variance was due to the integrators used. More than likely (and hopefully) most of this variability was due to the chromatography, although there is no way to determine this. Therefore, experimental analyses are not recommended to evaluate precision and accuracy of integrators.

Two remaining comparisons have been made. The first approach compares the peak integrator to manual peak-area determinations.[28,30] The second compares the response of one (proposed) integrator to a supposed standard, a commercial integrator.[3,4,20,27,132] Neither can be used reliably.

In the first case, planimetry[28] and triangulation[30] were used to demonstrate performance. As shown, neither manual method is very accurate nor precise, and comparisons to manual methods are not informative.

In the second case, many comparisons were made to point out restrictions in the older analog-integrator design, especially its horizontal baseline construction.[3,4,132] It was known that the analog integrator was less accurate in particular instances than the proposed device. For example, analog integrators could not tangent skim peaks, construct ramped baselines, nor recognize shoulder peaks. Additionally, their dynamic range was limited. Known standards were eluted under some of these conditions, and the two integrators were compared. Precision and relative accuracy were easily evaluated. In one other case,[20] comparison to a commercial integrator was implied to demonstrate absolute accuracy of the proposed software, similar to the traceability to NBS standards. This would be acceptable only after the commercial system itself was proven accurate.

D. Simulations

There are many advantages to using simulated chromatography:

1. There is more latitude in developing a systematic evaluation process.
2. This evaluation process will introduce the least error.
3. Properly modeled data can be representative of true chromatography.
4. It allows a systematic investigation on the effects of different types of noise.
5. The simulator can be brought to the data system.
6. Accuracy as well as precision can be investigated.
7. Evaluations can be time-efficient.

Critical to any meaningful simulation study are the models chosen for the peaks themselves

and the noise simulated. These have been discussed above. Also, consideration must be given to the simulation process itself; sampling rates, D/AC resolution, and precision timing are also important.

To date only two simulation studies have been published,[161,163] yet other unpublished studies exist.[74,164,165] Each has investigated different aspects of chromatography. The work of Hunt and Heisz[163] concerned peak height, retention time, and area differences at various levels of noise, overlap, and peak-height ratios. The work of Papas and Delaney[161] considered primarily the effects of noise, resolution deterioration, and tailing in isocratic chromatography. Unpublished work concerned overlapped-peak resolution, trace-peak quantitation, tangent skimmed peaks, and precision studies.

The study of Hunt and Heisz[163] simulated Gaussian chromatograms using a 16-bit A/DC at a 10-Hz sampling rate. Several aspects were investigated. The first was the effects of changing the relative height ratios and resolution in a group of three fused peaks of relative peak heights of 1:5:5. As predicted from the perpendicular drop method, increased negative bias was produced as the small peak either got smaller or became less resolved. No investigation of noise levels or tailing was performed. No comparison was made on their results and the theoretical values calculated from the standard error function (erf). These results reaffirm that fused peak-area determinations are done by the perpendicular drop method.

Tangent skimming effects was another area examined. Here one rider peak was varied in both relative area and resolution from the solvent peak. As predicted, an increasing larger negative bias was observed as the rider peak diminished in relative size. This bias stems from the way the peak is skimmed; instead of subtracting out a nonlinear segment of the underlying solvent peak, integrators construct a linear approximation. As the small peak becomes smaller, the relative error in using a linear vs. a nonlinear skimmed baseline becomes larger. This effect becomes magnified as the resolution of the smaller peak deteriorates.

In quantitating the effect of increased noise, the authors used a noise-generator chip. Noise was found to be substantial for both isolated peaks and unresolved peaks. Results for isolated peaks indicate a bias as large as -5% with a S/N of 20:1 increasing to -15% at a S/N of 5:1. This effect was more pronounced when overlapped peaks were overlaid with noise; a bias of $\pm 10\%$ was observed at a S/N of 20:1, whereas at 10:1, the bias was -10 to $+40\%$. These are exceedingly high values and do not agree with the other work.[161]

The authors verified the predicted inaccuracies of the perpendicular drop method and the linear baseline approximation used in tangent skimming. It might have been even more instructive if their results had been compared to the predicted bias values of the perpendicular drop peak-resolution method, easily calculated for Gaussian peaks. This would demonstrate subtle differences in where the perpendicular is dropped. The only result not consistent with other work was the substantial effect of noise. No explanation is offered, and these results may be due to problems with the simulation process itself.

There are two oversights within this study. The first problem is that a sampling rate of 10 Hz was used. To avoid aliasing problems as shown above, the simulation data should be sampled at a rate greater than that used by the system being tested. If this is not done, the variances observed will include the effect of aliasing; reproducibility of the simulation waveform relative to the A/DC sampling by the integrator is poor. This effect will also be seen by shifting retention times. The second problem concerns the use of Gaussian peaks. As discussed before, true chromatography is tailed and nonideal. Shown below, overlapped Gaussian peaks do not bear out difficulties seen with EMG peaks. Tailed peaks must be used to make the evaluation process both more authentic and more challenging.

The study by Papas and Delaney[161] focused on comparing several integrators for uniformity, bias, and precision. This included a limited study of tailed, noisy capillary peaks; a limited study with three forms of baseline disturbances; and a four-way experimental design study with four factors thought to effect peak accuracy.

Capillary peaks of 400-ms basewidth were simulated with one level of noise (S/N = 20) and four levels of tailing (tau/sigma = 0.5, 1.0, 1.5, 2.0). Full-scale peaks with the same areas were used throughout this study, yet no significant differences were found for the six integrators tested. Repeatability ranged from 1.04 to 0.033%. It is well known, that an inadequate sampling rate inhibits correct quantitation. Capillary peak simulation is also constrained by a suitable sampling rate; narrower peaks were not attempted to avoid aliasing problems, and wider peaks were used in another simulation test. Results here indicate that significant differences among integrators may only be due to sampling rate constraints. Indeed, the results of repeatability for 4-s peaks do not correspond to the 0.4-s capillary results. Therefore, accuracy for fast chromatography falls off at some point related to the sampling rate.

Three test chromatograms were simulated to test baseline construction. The first chromatogram consisted of a linear ramp baseline, seen with gradient profiles. The second chromatogram consisted of a broad, underlying peak, often seen with late eluting peaks or with severe drift. The last case involved a large, skewed peak representing a solvent peak. For each baseline test, the standard ten-peak chromatogram was used, with S/N = 6 and tau/sigma = 1.5 for each peak.

The ramped baseline results were uniformly consistent with those obtained with the linear baseline for all six integrators. The drifting baseline results were only uniform among the integrators tested and did not agree with the linear baseline results as expected. This was due to the valley-to-valley baseline construction used, since no commercial system can recognize an underlying peak. However, the tangent skim results were the most inconsistent in the study. There were discrepancies among integrators not only about the peak areas quantitated, but also on the number of peaks recognized. Additionally, some systems had to be forced to tangent skim the peaks. From this preliminary study, it appears this particular situation warrants more investigation. For laboratories heavily involved with tangent skimmed peaks, more rigorous simulation studies are necessary.

The final intraintegrator study consisted of a systematic investigation of noise, tailing, and peak width. Three levels of noise (S/N = 9, 6, 3), four levels of tailing (tau/sigma = 0.5, 1.0, 1.5, 2.0), and ten peak-width values (4, 8, 12, 16, 20, 24, 28, 32, 36, 40 s of basewidth) were used. The entire results have been published elsewhere,[16] so they will only be summarized here. Beyond what was predicted by the perpendicular drop method, two inconsistencies were noted in the results. The first concerned negative spikes, and the second concerned baseline construction for overlapped, tailed peaks.

As predicted from the perpendicular drop method for tailed, overlapped peaks, bias increased both as the level of tailing increased and as the resolution decreased. Noise did not appear to affect peak quantitation; neither did the peak width. However, when all ten tailed peaks overlapped into a fused group, two commercial integrators could not adaquately determine the baseline. Instead of dropping perpendiculars at valley points, baseline was forced through some of the valley points (see Figure 12). Varying the two peak parameters resulted only in forcing a larger number of erroneous baseline points or splitting two of the peaks; no correct parameters could be found that resulted in correct quantitation. The use of these particular units in this particular chromatographic situation only leads to erroneous results.

Negative spikes occurred when two or three simulated chromatograms were batched into one integrator file; between simulations, the D/AC dropped down to zero volts for several milliseconds. The algorithms for two systems tripped over this spike and insisted on calling it a baseline point. This resulted in large positive bias for second- and third-simulated injection (see Figure 13). With video-oriented integrators this problem is very obvious, whereas with only baseline codes for guidance, this problem can be almost impossible to detect.

This study simulated only a few representative chromatographic situations. Remaining

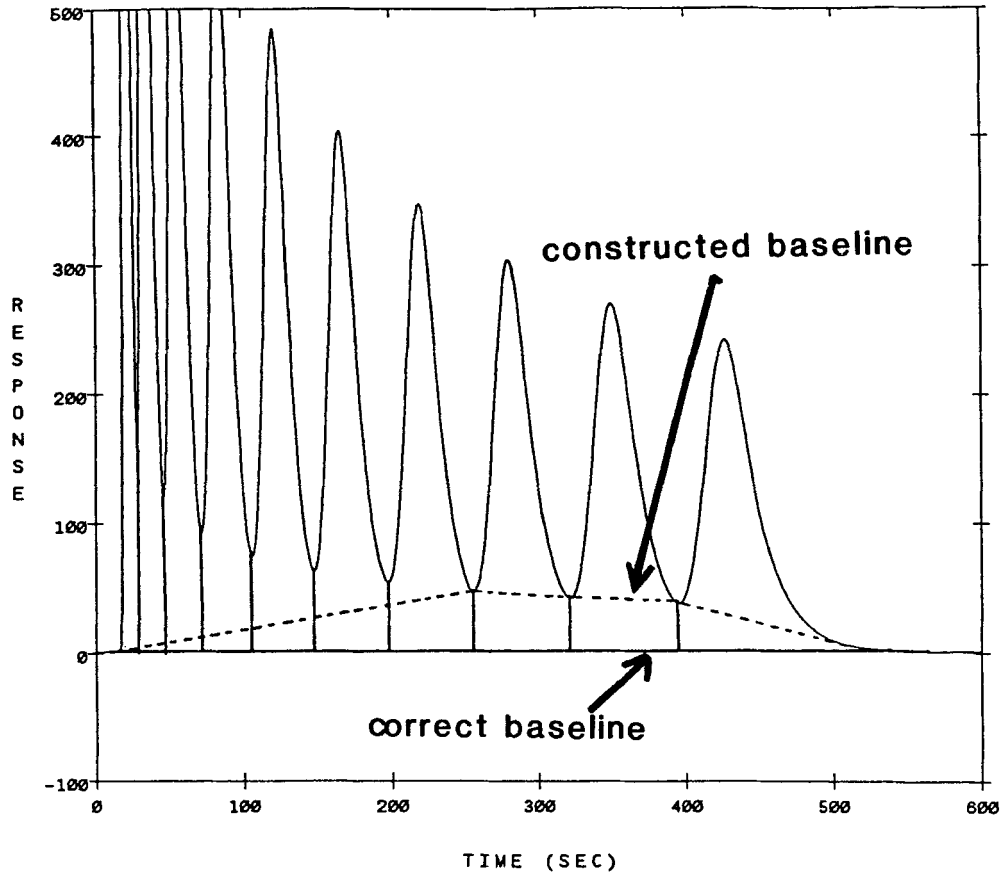


FIGURE 12. Incorrect baseline construction. Solid line is correct perpendicular drop, whereas dashed line is baseline construction observed with two chromatographic integrators.

are an infinite variety. Many systematic studies lay ahead that can be most effectively performed by simulation. As pointed out, Gaussian waveforms do not provide a rigorous test for integrator evaluation; results seen with tailed peaks do not correspond to the symmetrical ideal case. Caution must be given in selecting the peak and noise models, and in the choice of hardware. The operator must be certain that he or she is not introducing a bias into the evaluation process. When taking all these items into consideration, simulation studies are the most informative and efficient.

Both accuracy and precision studies are best performed by simulation, as evidenced by the levels observed with simulation studies vs. those obtained with chromatographed standard solutions. There is only one remaining concern before endorsing simulation as the only method necessary for QA testing. It is the age-old problem of representing all chromatographic data by one general function. To avoid oversimplification, it would seem necessary to include real chromatographic data. Simulation studies combined with experimental data that is stored and can be regenerated will become the most efficient manner to develop QA programs for integrator/data systems.

VI. CONCLUSIONS

This review has spanned over 30 years of development and advancement. As chromatography has moved from baseline-resolved full-scale peaks to multicomponent complex

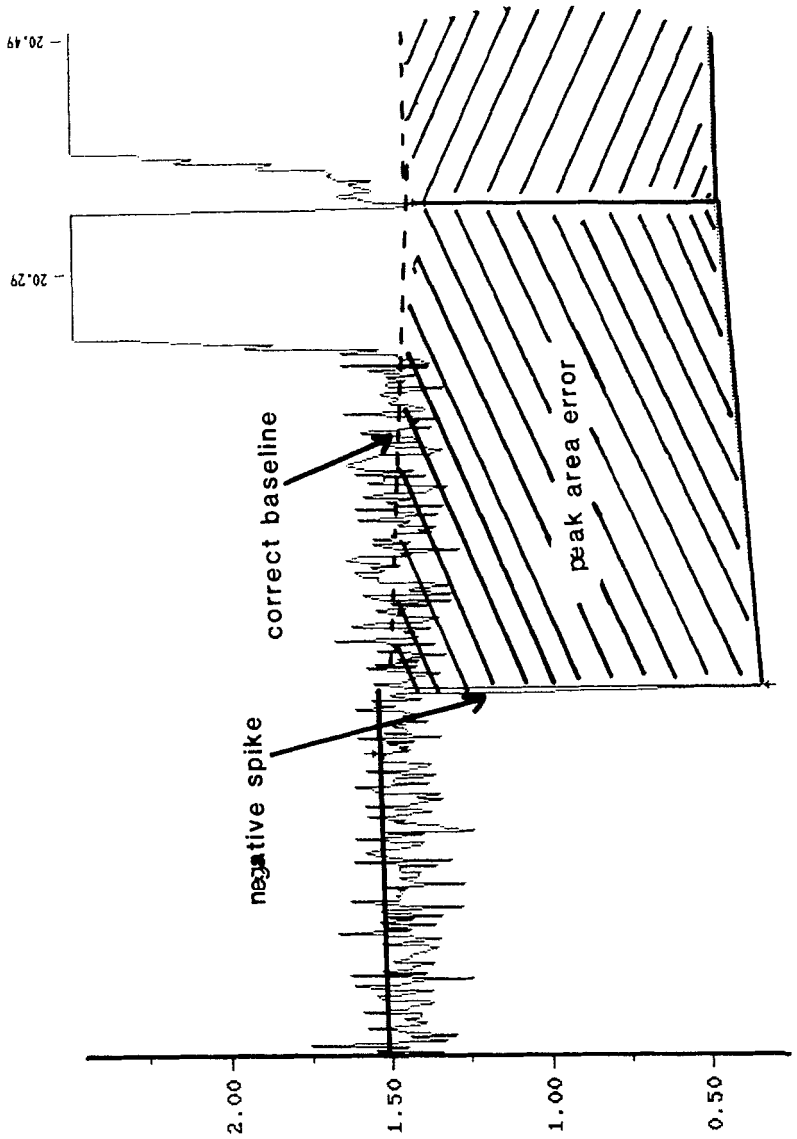


FIGURE 13. Negative spike between simulated injections that resulted in baseline beginning at bottom of spike.

environmental samples, rigorous demands have been placed on automatic peak-integrating devices. Although improvements have been made over the years to peak integrators for reliability and automation, the accuracy of overlapping peak deconvolution has not improved. Many alternatives have been tried, but commercial solutions remain unreliable. Present-day devices still perform peak resolution by an inherently inaccurate process used over 30 years ago, the perpendicular drop method. However, no new general-purpose peak-resolution technique has appeared on the horizon.

Even though integrators are black-box computer-based devices, they must still be treated as analytical instrumentation. Proper operation is not assured until the instrument responds well to known standards. QA programs for analytical instruments are a reality. In this paper, of the several approaches reviewed, simulation appears to be the most efficient method for systematically evaluating chromatographic integrators.

The opinions expressed in this review article are those of the author and are not necessarily those of the U.S. Food and Drug Administration. Reference to any commercial material, equipment, or process does not in any way constitute approval, endorsement, or recommendation by the U.S. Food and Drug Administration.

REFERENCES

1. **Gill, J. M.,** Automation of gas chromatography — concepts, characteristics and impact on future GC instrument design, *J. Chromatogr. Sci.*, 7, 731, 1969.
2. **Cotton, J. M.,** Approaches to laboratory data handling systems, *Am. Lab.*, 1, 41, 1969.
3. **Noonan, D. J. and March, E. W.,** Performance characteristics of a new GC data system, *Am. Lab.*, 20, 55, 1971.
4. **Baumann, F., Brown, A. C., and Mitchell, M. B.,** Chromatography and computer-non-routine GC laboratory applications, *J. Chromatogr. Sci.*, 8, 20, 1970.
5. **Healy, M. J. R., Chalmers, R. A., and Watts, W. E.,** Reduction of data from the automated gas-liquid chromatographic analysis of complex extracts from human biological fluids using a digital electronic integrator and an off-line computer programme, *J. Chromatogr.*, 87, 365, 1973.
6. **Wilson, W. O. and Price, J. G. W.,** Installation and operation of similar gas chromatography-computer systems in R/D and quality control laboratories, *J. Chromatogr. Sci.*, 8, 31, 1970.
7. **Hettinger, J. D., Hubbard, J. R., Gill, J. M., and Miller, L. A.,** A new computing integrator for chromatography, *J. Chromatogr. Sci.*, 9, 711, 1971.
8. **Giddings, J. C.,** *Dynamics of Chromatography, Part I: Principles and Theory,* Marcel Dekker, New York, 1965, 20.
9. **Giddings, J. C.,** *Gas Chromatography, 1964,* Goldup, A., Ed., Elsevier, Amsterdam, 1964.
10. **Vink, H.,** Theory of partition chromatography. III, *J. Chromatogr.* 20, 305, 1965.
11. **Snyder, L. R. and Kirkland, J. J.,** Quantitative and trace analysis, in *Introduction to Modern Liquid Chromatography,* 2nd ed., John Wiley & Sons, New York, 1979, 556.
12. **Janik, A.,** What is the best measure of the detector signal: the peak area or the peak height?, *J. Chromatogr. Sci.*, 13, 93, 1975.
13. **Bakalayer, R. and Henry, S.,** Variables affecting precision and accuracy in high performance liquid chromatography, *J. Chromatogr.*, 126, 327, 1976.
14. **Ball, D. L., Harris, W. E., and Habgood, H. W.,** Experimental evaluation of indeterminate error in height-width and height only measurements of chromatographic peaks, *Anal. Chem.*, 40, 129, 1968.
15. Subcommittee E-19.08 Task Group on Liquid Chromatography for the ASTM, An evaluation of quantitative precision in high performance liquid chromatography, *J. Chromatogr. Sci.*, 19, 338, 1981.
16. **Pauls, R. E., McCoy, R. W., Ziegler, E. R., Wolf, T., Fritz, G. T., and Marmion, D. M.,** Results of a cooperative study comparing the precision and accuracy of peak height and area measurements in liquid chromatography. II, *J. Chromatogr. Sci.*, 24, 273, 1986.
17. **Kipiniak, W.,** A basic problem — the measurement of height and area, *J. Chromatogr. Sci.*, 19, 332, 1981.
18. **Bobba, G. M. and Donaghey, L. F.,** Real time analysis of gas chromatography with a microcomputer system, *J. Chromatogr. Sci.*, 15, 47, 1977.

19. **Metzger, H. D. and Radszurveit, K. G.,** Program and system for the evaluation of gas chromatographic analysis, *Chromatographia*, 5, 186, 1972.
20. **Sickenberger, D. W.,** Data acquisition and control system, *Am. Lab.*, 20, 70, 1984.
21. **Tse, R. S.,** An inexpensive chromatographic data system for the Apple II+, *Comput. Chem.*, 9, 317, 1985.
22. **Smith, R. J. and Lloyd, L. G.,** Microcomputer integration of chromatographic data, *IRCS Med. Sci. Biochem. Biomed. Technol.*, 10, 117, 1982.
23. **Khan, S. A.,** Chromatographic data acquisition and data analysis system . I, *Instrum. Comput.*, 12, January 1986.
24. **Burke, M. F. and Thurman, R. G.,** Use of a dedicated computer for real-time control of gas chromatographic measurement, *J. Chromatogr. Sci.*, 8, 39, 1970.
25. **Shank, J. T. and Persinger, H. E.,** A precision measuring system for gas chromatography, *J. Gas Chromatogr.*, 5, 631, 1967.
26. **Grisham, C. M., Marquard, F., and Jorgensen, P. L.,** Versatile high performance liquid chromatographic computer system for solvent delivery, gradient control, and data acquisition using the Commodore 64 personal computer, *J. Chromatogr.*, 333, 301, 1985.
27. **Lyne, P. M. and Scott, K. F.,** Commodore 3032 computer controlled chromatography for physicochemical and analytical applications. II. Simple peak integration, kinetics by stopped-flow chromatography and automatic adsorption isotherms, *J. Chromatogr.*, 19, 599, 1981.
28. **Hegedus, L. L. and Petersen, E. E.,** Integration of chromatographic signals by digital computers: an approach for the small chromatographic laboratory where digital computer services are available, *J. Chromatogr. Sci.*, 9, 551, 1971.
29. **Giles, I. G. and Gore, M. G.,** Biochemical data processing with microcomputers. II. A basic program for the detection, integration, and area assignment of chromatographic elution profiles, *Anal. Chim. Acta*, 151, 123, 1983.
30. **Overton, M. W., Alber, L. L., and Smith, D. E.,** Time-sharing minicomputer data acquisition-processing system, *Anal. Chem.*, 47, 363A, 1975.
31. **Grushka, E., Myers, M. N., Schettler, P. D., and Giddings, J. C.,** Computer characterization of chromatographic peaks by peak height and higher central moments, *Anal. Chem.*, 41, 889, 1969.
32. **Knoll, J. E. and Midgett, M. R.,** The area overlap method for determining adequate chromatographic resolution, *J. Chromatogr. Sci.*, 20, 221, 1982.
33. **Kirkland, J. J., Yau, W. W., Stoklosa, H. J., and Diiks, C. H.,** Sampling and extra-column effects in high performance liquid chromatography; influence of peak skew on plate count calculations, *J. Chromatogr. Sci.*, 15, 303, 1977.
34. **Chesler, S. P. and Cram, S. P.,** Effect of peak sensing and random noise on the precision and accuracy of statistical moment analyses from digital chromatographic data, *Anal. Chem.*, 43, 1922, 1971.
35. **Yau, W. W.,** Characterizing skewed chromatographic band broadening, *Anal. Chem.*, 49, 395, 1977.
36. **Anderson, D. J. and Walters, R. R.,** Effect of baseline errors on the calculation of statistical moments of tailed chromatographic peaks, *J. Chromatogr. Sci.*, 22, 353, 1984.
37. **Yamoaka, K. and Nakagawa, T.,** Moment analysis of non-linear gas chromatography by computer simulation, *J. Chromatogr.*, 103, 221, 1975.
38. **Grushka, E., Myers, M. N., and Giddings, J. C.,** Moment analysis for the discernment of overlapping chromatographic peaks, *Anal. Chem.*, 42, 21, 1970.
39. **Jung, K. H., Yun, S. J., and Kang, S. H.,** Characterization of peak shape parameters with normal and derivative chromatograms, *Anal. Chem.*, 56, 457, 1984.
40. **Grushka, E.,** Characterization of exponentially modified Gaussian peaks in chromatography, *Anal. Chem.*, 44, 1733, 1972.
41. **Grubner, O.,** Interpretation of asymmetric curves in linear chromatography, *Anal. Chem.*, 43, 1934, 1971.
42. **Foley, J. P. and Dorsey, J. G.,** A review of the exponentially modified gaussian (EMG): evaluation and subsequent calculation of universal data, *J. Chromatogr. Sci.*, 22, 1984.
43. **Foley, J. P. and Dorsey, J. G.,** Equations for calculation of chromatographic figures of merit for ideal and skewed peaks, *Anal. Chem.*, 55, 730, 1983.
44. **Foley, J. P.,** Equations for chromatographic peak modeling and calculation of peak area, *Anal. Chem.*, 59, 1984, 1987.
45. **Barber, W. E. and Carr, P. W.,** Graphical methods for obtaining retention time and number of theoretical plates from tailed chromatographic peaks, *Anal. Chem.*, 53, 1939, 1981.
46. **Hanggi, D. and Carr, P. W.,** Errors in exponentially modified Gaussian equations in the literature, *Anal. Chem.*, 57, 2394, 1985.
47. **Crispin, T. and Wirtz, R.,** Qualification and validation of hplc systems with CR3a integrator. A BASIC program for direct calculation of important chromatographic data, *Chromatographia*, 22, 434, 1986.
48. **Giddings, J. C.,** Kinetic origin of tailing in chromatography, *Anal. Chem.*, 35, 1999, 1963.

49. **Kucera, E.,** Contribution to the theory of chromatography. Linear non-equilibrium elution chromatography, *J. Chromatogr.*, 19, 235, 1965.
50. **McWilliams, I. G. and Bolton, H. C.,** Instrumental peak distortion. II. Effect of recorder response time, *Anal. Chem.*, 41, 1762, 1969.
51. **Buys, T. S. and de Clerk, K.,** Bi-gaussian fitting of skewed peaks, *Anal. Chem.*, 44, 1274, 1972.
52. **Jonsson, J. A.,** Computer simulation of the chromatographic process, *Chromatographia*, 13, 273, 1980.
53. **Fraser, R. D. B. and Suzuki, E.,** Resolution of overlapping absorption bands by least squares procedures, *Anal. Chem.*, 38, 1770, 1966.
54. **Fraser, R. D. B. and Suzuki, E.,** Resolution of overlapping bands: functions for simulating band shapes, *Anal. Chem.*, 41, 37, 1969.
55. **Chesler, S. N. and Cram, S. P.,** Iterative curve fitting of chromatographic peaks, *Anal. Chem.*, 45, 1354, 1973.
56. **Pauls, R. E. and Rogers, L. B.,** Band broadening studies using parameters for an exponentially modified gaussian, *Anal. Chem.*, 49, 625, 1977.
57. **Maynard, V. and Grushka, E.,** Effect of dead volume on efficiency of a gas chromatographic system, *Anal. Chem.*, 44, 1427, 1972.
58. **Binsheng, L., Haochun, L., and Peichang, L.,** A computer-aided algorithm for saving "living" chromatograms in a knowledge base, *LC-GC Mag.*, 4, 1206, 1986.
59. **Anderson, A. H., Gibb, T. C., and Littlewood, A. B.,** Computer resolution of unresolved convoluted gas-chromatographic peaks, *J. Chromatogr. Sci.*, 8, 640, 1970.
60. **Mott, S. D. and Grushka, E.,** Chromatographic solute identification using peak shape analysis, *J. Chromatogr.*, 126, 191, 1976.
61. **Li, K. P. and Li, Y. Y. H.,** Rate dependence of statistical moments of chromatographic profile on solvent-solvent interactions, *Anal. Chem.*, 48, 77, 1976.
62. **Lundeen, J. T. and Juvet, R. S.,** Quantitative resolution of severely overlapped chromatographic peaks, *Anal. Chem.*, 53, 1369, 1981.
63. **Smit, J. C., Smit, H. C., and DeJager, E. M.,** Computer implementation of simulation models for non-linear, non-ideal chromatography. I. Fundamental mathematical considerations, *Anal. Chim. Acta*, 122, 1, 1980.
64. **Smit, J. C., Smit, H. C., and DeJager, E. M.,** Computer implementation of simulation models for non-linear, non-ideal chromatography. II. Numerical experiments and results, *Anal. Chim. Acta*, 122, 151, 1980.
65. **Ng, T. L. and Ng, S.,** Computer simulation of abnormal high-performance liquid chromatograms caused by solvents, *J. Chromatogr.*, 329, 13, 1985.
66. **Novak, J., Petrovic, K., and Wicar, S.,** Efficacy of the corrections applied to the resolution of overlapping chromatographic peaks by the perpendicular drop method, *J. Chromatogr.*, 55, 221, 1971.
67. **Westerberg, A. W.,** Detection and resolution of overlapping peaks for an on-line computer system for gas chromatographs, *Anal. Chem.*, 41, 1770, 1969.
68. **Hancock, H. A., Dahm, L. E., and Muldoon, J. F.,** A new computer program for the resolution of fused chromatographic peaks, *J. Chromatogr. Sci.*, 8, 57, 1970.
69. **Proksch, E., Bruneder, H., and Granzer, V.,** Correction factors for the quantitative evaluation of overlapping gas chromatographic peaks, *J. Chromatogr. Sci.*, 7, 473, 1969.
70. **Landowne, R. A., Morosani, R. W., Herrmann, R. A., King, R. M., and Schmus, H. G.,** Computer acquisition and analysis of gas chromatographic data, *Anal. Chem.*, 44, 1961, 1972.
71. **Ball, D. L., Harris, W. E., and Habgood, H. W.,** Errors in manual integration techniques for chromatographic peaks, *J. Gas Chromatogr.*, 5, 613, 1967.
72. **Mori, Y.,** A practical quantitative method for unresolving gas chromatographic peaks, *J. Chromatogr.*, 66, 9, 1972.
73. **Delaney, M. F.,** personal communication, 1986.
74. **Chidester, D.,** personal correspondence, American Cyanamid, 1987.
75. **Papas, A.,** unpublished data, 1986.
76. Peak-11 manual, PDP computer systems, Digital Equipment Corp., Maynard, MA.
77. **O'Haver, T. C. and Green, G. L.,** Numerical error analysis of derivative spectrometry for the quantitative analysis of mixtures, *Anal. Chem.*, 48, 312, 1976.
78. **Shukla, S. S. and Rusling, J. F.,** Comparison of methods for numerical differentiation, *Trends Anal. Chem.*, 4, 229, 1985.
79. **Savitzky, A. and Golay, M. J. E.,** Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem.*, 36, 1627, 1964.
80. **Kambara, T. and Saitoh, K.,** Derivative gas chromatography applied to the micro-determination of the poorly resolved component, *J. Chromatogr.*, 35, 318, 1968.
81. **Ashley, J. W. and Reilley, C. N.,** De-tailing and sharpening of response peaks in gas chromatography, *Anal. Chem.*, 37, 626, 1965.

82. **Grushka, E. and Monacelli, G. C.,** Slope analysis for recognition and characterization of strongly overlapped chromatographic peaks, *Anal. Chem.*, 44, 484, 1972.
83. **Butterfield, R. O., Lancaster, E. B., and Dutton, H. J.,** Analog computer program for resolution of overlapping distribution curves, *Sep. Sci.*, 1, 329, 1966.
84. **Anderson, A. H., Gibb, T. C., and Littlewood, A. B.,** Computer analysis of unresolved non-Gaussian gas chromatography by curve-fitting, *Anal. Chem.*, 42, 434, 1970.
85. **Gladney, H. M., Dowden, B. F., and Swalen, J. D.,** Computer-assisted gas-liquid chromatography, *Anal. Chem.*, 41, 883, 1969.
86. **Roberts, S. M.,** Short cut fused peak resolution methods for chromatograms, *Anal. Chem.*, 44, 502, 1972.
87. **Hippe, Z., Bierowska, A., and Pietryga, T.,** Algorithms for high-level data processing in gas chromatography, *Anal. Chim. Acta*, 122, 279, 1980.
88. **Goldberg, B.,** A numerical method of resolving peak areas in gas chromatography, *J. Chromatogr. Sci.*, 9, 287, 1971.
89. **Gourlia, J. P. and Bordet, J.,** Fast automatic gas chromatographic analysis with short columns, *J. Chromatogr. Sci.*, 19, 35, 1981.
90. **Schreen, P. J. H., Klous, Z., Smit, H. C., and Doornbos, D. A.,** A software package for the orthogonal polynomial approximation of analytical signals, including a simulation program for chromatograms and spectra, *Anal. Chim. Acta*, 171, 45, 1985.
91. **Macnaughtan, D., Rogers, L. B., and Wernimont, G.,** Principal component analysis applied to chromatographic data, *Anal. Chem.*, 44, 1421, 1972.
92. **Woodruff, H. B., Tway, P. C., and Love, L. J. C.,** Factor analysis of mass spectra from partially resolved chromatographic peaks using simulated data, *Anal. Chem.*, 53, 81, 1981.
93. **Frank, I. E. and Kowalski, B. R.,** Chemometrics review, *Anal. Chem.*, 54, 232R, 1982.
94. **Delaney, M. F.,** Chemometrics review, *Anal. Chem.*, 56, 128A, 1984.
95. **Kowalski, B. R. et al.,** Chemometrics review, *Anal. Chem.*, 58, 240A, 1986.
96. **Mauro, D. M. and Delaney, M. F.,** Resolution of infared spectra of mixtures by self-modeling curve resolution using a library of reference spectra with simplex-assisted searching, *Anal. Chem.*, 58, 2622, 1986.
97. **Ramos, L. S., Burger, J. E., and Kowalski, B. R.,** Third-order chromatography: multivariate analysis of data from parallel-column chromatography with multichannel detection, *Anal. Chem.*, 57, 2620, 1985.
98. **Hieftje, G. M.,** Signal-to-noise enhancement through instrumental techniques. I. Signals, noise, and s/n enhancement in the frequency domain, *Anal. Chem.*, 44, 81A, 1972.
99. **Hieftje, G. M.,** Signal-to-noise enhancement through instrumental techniques. II. Signal averaging, boxcar integration, and correlation techniques, *Anal. Chem.*, 44, 69A, 1972.
100. **Goedert, M. and Guiochon, G.,** Sources of error in chromatographic analysis. Effect of sampling parameters on the accuracy of numerical measurements, *Chromatographia*, 6, 76, 1973.
101. **Goedert, M. and Guiochon, G.,** Effect of signal noise on the determination of retention time of peak maximum. Computer simulation and experimental study, *Chromatographia*, 6, 39, 1973.
102. **Enke, C. G. and Nieman, T. A.,** Signal-to-noise ratio enhancement by least-squares polynomial smoothing, *Anal. Chem.*, 48, 705A, 1976.
103. **Thompson, M. R. and Dessey, R. E.,** Use and abuse of digital signal processors, *Anal. Chem.*, 56, 583, 1984.
104. **Cram, S. T., Chesler, S. N., and Brown, A. C.,** Effects of digital filters on chromatographic signals, *J. Chromatogr.*, 126, 279, 1976.
105. **Nevius, T. A. and Pardue, H. L.,** Development and preliminary evaluation of modified Savitzky-Golay smoothing functions, *Anal. Chem.*, 56, 2249, 1984.
106. **Bromba, M. U. and Ziegler, H.,** Digital filter for computationally efficient smoothing of noisy spectra, *Anal. Chem.*, 55, 1299, 1983.
107. **Bromba, M. U. and Ziegler, H.,** Variable filter for digital smoothing and resolution enhancement of noisy spectra, *Anal. Chem.*, 56, 2052, 1984.
108. **Nickolaisen, S. L. and Bialkowski, S. E.,** A least-squares filter for repetitive data acquisition, *J. Chem. Inf. Comput. Sci.*, 26, 57, 1986.
109. **Malczewski, M. L. and Grushka, E.,** Multiple peak recognition in high performance liquid chromatography by fast Fourier transformation, *J. Chromatogr. Sci.*, 19, 187, 1981.
110. **Kirmse, D. W. and Westerberg, A. W.,** Resolution enhancement of chromatographic peaks, *Anal. Chem.*, 43, 1035, 1971.
111. **Miscoski, M. F.,** The HP3393A computing integrator, *Chromatography*, 1, 49, 1987.
112. **Laeven, J. M., Smit, H. C., and Lankelma, J. V.,** A software package for the generation of noise with widely divergent spectral properties, *Anal. Chim. Acta*, 157, 373, 1984.
113. **Synovac, R. E. and Yeung, E. S.,** Improvement of the limit of detection in chromatography by an integration method, *Anal. Chem.*, 57, 2162, 1985.

114. **de Root, G. J.,** Fast Fourier transformation filter program as part of a system for processing chromatographic data with an Apple II microcomputer, *Trends Anal. Chem.*, 4, 134, 1985.
115. **Shukla, S. S. and Rusling, J. F.,** Analyzing chemical data with computers: errors and pitfalls, *Anal. Chem.*, 56, 1347A, 1984.
116. **Dessey, R. E.,** The electronic toolbox. II, *Anal. Chem.*, 57, 1298A, 1985.
117. **Johnson, D., Lawson, D. D., and Havlik, A. J.,** Voltage-to-frequency integrators in gas chromatography, *J. Gas Chromatogr.*, 5, 303, 1965.
118. **Johnson, H. W.,** Storage and complete automatic computation of gas chromatographic data, *Anal. Chem.*, 35, 521, 1963.
119. **Sawyer, D. T. and Barr, J. K.,** Evaluation of several integrators for use in gas chromatography, *Anal. Chem.*, 34, 1213, 1962.
120. **Derge, K.,** Eine Ubersicht der bestehenden Systeme fur die Datenerfassung in Chromatographie-Labor, *J. Chromatogr.*, 77, 75, 1973.
121. **Baumann, F. and Tao, F.,** Digital integrators-effect of slope sensitivity, filtering, and baseline correction rate on accuracy, *J. Gas Chromatogr.*, 5, 621, 1967.
122. **Karohl, J. G.,** Peak detection logic systems of gas chromatographic integrators, *J. Gas Chromatogr.*, 5, 627, 1967.
123. **Janak, J.,** Statistische Auswertung der Messmethoden zur registrierten Differentialkurven in der Gas-Chromatographie, *J. Chromatogr.*, 3, 308, 1960.
124. **Delaney, M. F.,** Systematic errors in manual area measurement of skewed chromatographic peaks, *Analyst*, 107, 606, 1982.
125. **Bohm, Z.,** Ein Analogintegrator zur gaschromatographischen Apparutur, *J. Chromatogr.*, 3, 265, 1960.
126. **Strickler, A. and Gallaway, W. S.,** Analog integration techniques in chromatographic analysis, *J. Chromatogr.*, 5, 185, 1961.
127. **Ouchi, G. I.,** Use your lab pc to handle all of your chromatography data, *Res. Dev.*, 26, 150, 1984.
128. **Rooney, T.,** Increasing productivity in the chromatography laboratory, *Am. Lab.*, 20, 55, 1983.
129. **Blow, A. M. J. and Langford, B.,** A new chromatography data system, *Am. Lab.*, 20, 60, 1983.
130. **Rooney, T.,** A networked computer architecture for enhanced chromatography productivity, *Am. Lab.*, 21, 58, 1987.
131. **Sato, T., Takimoto, I., Fujii, S., and O'Neill, O.,** An economical high performance chromatographic data processor, *Am. Lab.*, 21, 122, 1985.
132. **Bilen, L. and Mikkelsen, L.,** Quantitative performance of the HP 3357 for capillary column peaks, *Lab. Press Rev.*, 2, 30, 1983.
133. **Rooney, T.,** Power and economy in a new computing integrator, *Lab. Press Rev.*, 2, 37, 1983.
134. **Andrews, M.,** The Waters 820 chromatography workstation, *Chromatographia*, 2, 21, 1987.
135. **Bruce, G. L.,** The Varian ds-600 multi-channel chromatography data system, *Chromatographia*, 2, 29, 1987.
136. **Rooney, T.,** ChromaStation: the chromatography link to Lotus from Spectra-Physics, *Chromatographia*, 2, 43, 1987.
137. **Andrews, M. and Cotter, R.,** Computerized control and data handling in the hplc laboratory, *Am. Lab.*, 20, 32, 1984.
138. **Day, J. A.,** Evaluation of the Perkin-Elmer sigma 10 data station and comparison with a centralised mini-computer, *Analyst*, 697, 1980.
139. **Tochner, M., Magnuson, J. A., and Soderman, L. Z.,** Time-sharing — a powerful approach for gas chromatographic data reduction, *J. Chromatogr. Sci.*, 7, 740, 1969.
140. **Shaw, P. E. and Miller, J. M.,** Interfacing an Apple computer with a digital integrator for chromatographic data storage and statistical calculations, *J. Chromatogr. Sci.*, 21, 372, 1983.
141. **Tivin, F.,** Routine and research compatibility in computer assisted gas chromatography, *J. Chromatogr. Sci.*, 8, 13, 1970.
142. **Raymond, A. J., Lawrey, D. M. G., and Mayer, T. J.,** Acquisition and processing of gas chromatographic data using a time-shared computer, *J. Chromatogr. Sci.*, 8, 1, 1970.
143. **McCullough, R. D.,** Gas chromatograph data handling with on-line data acquisition computers, *J. Gas Chromatogr.*, 5, 635, 1967.
144. **Landowne, A., Morosani, R. W., Herrmann, R. A., King, R. M., and Schmus, H. G.,** Computer acquisition and analysis of gas chromatographic data, *Anal. Chem.*, 44, 1961, 1972.
145. **Ziegler, E. and Schomburg, G.,** The application of COLACHROM, a new command language for chromatographic data processing, *J. Chromatogr.*, 290, 339, 1984.
146. **Reese, C. E.,** Chromatographic data acquisition and processing. II. Data manipulation, *J. Chromatogr. Sci.*, 18, 249, 1980.
147. **Woerlee, E. F. G. and Mol, J. C.,** A real-time gas chromatographic data system for laboratory application, *J. Chromatogr. Sci.*, 18, 258, 1980.

148. **Anderson, R. E.,** Trends in chromatographic automation, *Chromatographia*, 5, 105, 1972.
149. **Baumann, F., Herlicska, E., and Brown, A. C.,** Quantitative evaluation of chromatography by digital computers, *J. Chromatogr. Sci.*, 7, 680, 1969.
150. **Leung, A. T., Rock, J. V., Henry, R. A., Derge, K., and McIllwrick, R.,** Effect of microprocessors on traditional chromatographic instrumentation, *J. Chromatogr.*, 122, 355, 1976.
151. **Emery, E. M.,** The role of new generation chromatography automation in industrial research, *J. Chromatogr. Sci.*, 14, 261, 1976.
152. Guide to inspection of computerized systems in drug processing, *Pharm. Manuf.*, March 39, 1984.
153. **Kuzel, N. R.,** Quality assurance auditing of computer systems, *Pharm. Technol.*, 34, February, 1987.
154. Software quality assurance plans, ANSI/IEEE Standard No. 730-1984.
155. **Dunn, R. and Ullman, R.,** *Quality Assurance for Computer Software*, McGraw-Hill, New York, 1982.
156. Software quality evaluation, DOD Standard No. DOD-STD-2168.
157. Software Development Activities, Tech. Rep. U.S. Food and Drug Administration, July, 1987.
158. Quality Control Reports, *Gold Sheets*, 21, 6, 1987.
159. **Dolan, J. W.,** LC user survey: data system problems, *LC-GC Mag.*, 4, 620, 1986.
160. Assistant editor, Managing chromatographic data. Focus, *Anal. Chem.*, 57, 1256A, 1985.
161. **Papas, A. N. and Delaney, M. F.,** Evaluation of chromatographic integrators and data systems, *Anal. Chem.*, 59, 54A, 1987.
162. **Johnson, R. D.,** A peak generator for gas chromatography, *J. Gas Chromatogr.*, 6, 43, 1968.
163. **Hunt, R. J. and Heisz, K.,** Evaluation of results of data processing systems, *GIT Fachz. Lab.*, 30, 789, 1986.
164. **Newcomb, R.,** personal correspondence, Kodak, 1987.
165. **Eglington, D.,** unpublished data, 1987.
166. **Papas, A. N.,** unpublished data, 1981.