Journal of Chromatography, 384 (1987) 301–313
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1018

SYSTEMATIC ERRORS IN THE MEASUREMENT OF PEAK AREA AND PEAK HEIGHT FOR OVERLAPPING PEAKS

JOE P. FOLEY
Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803-1804 (U.S.A.)

SUMMARY

Errors in the algorithms commonly employed by integrators and data systems for the measurement of peak areas and peak heights of overlapped peaks are re-evaluated for cases where one or both peaks is stailed. The errors due only to peak tailing (area ratio = 1) can be greater than 50% for peak area measurement by the perpendicular-drop algorithm; errors in peak height can be greater than 40%. Errors due to a combination of peak tailing and differences in peak size can exceed 200% for peak area and 80% for peak height. An empirical area equation, when used in conjunction with normal integration procedures, permits the accurate (< ±4%) quantitation of overlapping peaks, provided that the valley between the peaks is less than 45%.

INTRODUCTION

Gas and liquid chromatography are powerful methods for the quantitative analysis of multi-component mixtures. Unfortunately, as the number of components in a mixture increases, the probability that all components will be baseline resolved drops precipitously[1,2]. Inevitably at least some of the partially resolved peaks will be of interest to the analyst. Thus, the quantitation of overlapping peaks is an important issue.

The quantitation of chromatographic peaks requires the measurement of either peak height or peak area. While the manual or electronic measurement of peak heights and peak areas for baseline-resolved peaks is straightforward and accurate, the same measurement for partially resolved peaks is neither, due to the distortion caused by peak overlap. Although peak deconvolution methods have been developed[3–7], the approach used by nearly all electronic integrators and most data systems (and the only approach available when manual measurements are made) is to approximate the peak areas and peak heights of overlapped peaks by measurements made directly from the overlapping chromatogram. For peak areas, the perpendicular-drop and tangent-skimming methods are used on overlapping peaks with comparable and disproportionate areas, respectively. For peak height, the apparent peak height is used in all cases.

The accuracy of these integrator–data system approximations for overlapping

0021-9673/87/$03.50

peaks has been evaluated extensively[8–10]. It was found that peak height could be measured much more accurately than peak area for a given resolution and peak area ratio. Many refinements have been proposed, particularly for peak area measurements[8], including the use of correction factors based on calculated errors[11,12].

Unfortunately, although these overlapping peaks approximations have been evaluated in detail, previous researchers have assumed symmetric (Gaussian) peak profiles for the component peaks. However, it is clear from the chromatographic literature that many overlapping peak profiles frequently consist of asymmetric (tailed) chromatographic peaks. Our goals in the present study are (1) to re-evaluate the accuracy of these popular approximations for the measurement of peak area and peak height for overlapping peaks when one or both peaks are tailed; and (2) to investigate alternative approaches for the quantitation of overlapping peaks that do not require deconvolution. Because our emphasis is on the effect of peak asymmetry on overlapping peaks and not on the effect of relative peak size, which has already been extensively discussed (for overlapping Gaussian peaks)[9,10], we limited our studies to overlapping peaks with area ratios of 4 or less. Thus our present study does not include a re-evaluation of the tangent skimming method, since this method is inappropriate under these circumstances.

EXPERIMENTAL

Computations. An Apple Macintosh computer programmed in BASIC was used for all calculations.

Symmetric and asymmetric (tailed) peak models. Symmetric chromatographic peaks were generated using a normalized Gaussian function $G(t)$,

$$G(t) = A/\sigma_G(2\pi)^{\frac{1}{2}}] \exp[-(t-t_G)^2/2\sigma_G{}^2] \qquad (1)$$

where A is the peak area, t_G is the retention time, and σ_G is the standard deviation of the peak. Tailed chromatographic peaks were generated using a normalized exponentially modified Gaussian function $EMG(t)$. The EMG function results from the convolution of a Gaussian function and an exponential decay function and can be expressed in a variety of ways[13–15]. The specific form we used was

$$EMG(t) = A/\tau \exp[1/2(\sigma_G/\tau)^2 - (t-t_G)/\tau] \int_{-\infty}^{z} \exp(-y^2/2)/(2\pi)^{\frac{1}{2}} dy \qquad (2)$$

where A is the peak area, t_G and σ_G are the retention time and standard deviation of the Gaussian function respectively, τ is the time constant from the exponential decay function, and $z = (t-t_G)/\sigma_G - \sigma_G/\tau$. The integral in eqn. 2 was evaluated as previously described[16]. Note that the ratio τ/σ_G is a fundamental measure of peak symmetry. As τ/σ_G increases, the tailing of the chromatographic peak increases. As τ/σ_G approaches 0, the resulting peak approaches that of a Gaussian.

Peak overlap simulations. A pair of overlapping peaks was simulated by adding the functions representing the individual chromatographic peaks. Four combinations of overlapping peaks (symmetric–symmetric, tailed–tailed, tailed–symmetric, and symmetric–tailed) were examined for five area ratios (0.25, 0.5, 1, 2, and 4), five τ/σ_G ratios (0.5, 1, 2, 3, and 4), and several values of resolution, resulting in over 600 pairs of overlapping peaks. Resolution was defined conventionally as $\Delta t_G/4$ (variance)$^{\frac{1}{2}}$, where $\Delta t_G = t_{G,2} - t_{G,1}$ and σ_G^2 and $\sigma_G^2 + \tau^2$ are the variance of a Gaussian and an EMG peak, respectively. A constant Gaussian contribution to the total variance (fixed value of σ_G) was assumed for both peaks in every peak combination, although the EMG–EMG peak combinations were also examined from the point of view of constant total variance. Resolution values of 1.75, 1.5, 1.25, 1.125, 1, 0.875, 0.625, 0.5, 0.375, and 0.25 were used in generating the data for this study. Although the resolution parameter proved useful in generating the data, it was found to be inadequate in describing peak overlap in cases where one or both peaks is tailed (see Results and discussion).

Measurement of peak parameters. A search algorithm for the measurement of the retention time, peak height, and width and asymmetry of an isolated peak at any peak height fraction[16] was modified to measure the pertinent parameters of overlapping peaks (see Fig. 2). in simulating the perpendicular-drop algorithm, peak areas were calculated by summation from the valley to the appropriate baseline. This is the method most commonly employed by chromatographic integrators and/or data systems.

RESULTS AND DISCUSSION

Preliminary considerations

Peak modeling. The Gaussian and the exponentially modified Gaussian (EMG) functions were used as models for symmetric and asymmetric (tailed) chromatographic peaks. Four combinations of overlapping peaks, described below, were examined in our evaluation of the perpendicular-drop and apparent-peak-height methods for peak area and peak height measurement.

(1) Symmetric peak–symmetric peak. Neither peak is subject to asymmetric band-broadening processes. This combination has been the only one examined in nearly all prior studies of overlapping peaks.

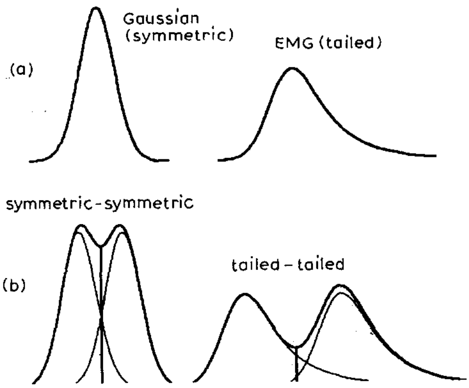
(2) Tailed peak–tailed peak. Both peaks are subject to asymmetric band-broadening processes. This is most likely to be observed when extra-column effects (which affect closely eluted peaks similarly) cannot be eliminated, although it could also be observed if both peaks participated in an irreversible retention mechanism.

(3) Tailed peak–symmetric peak. Only the first peak is subject to an asymmetric band-broadening process. This could occur when an additional, irreversible (slow) retention mechanism is operative for one of the peaks.

(4) Symmetric peak–tailed peak. Same as 3, except that the tailed peak is eluted last.

To perform the simulations, values for the variances of the overlapping peaks must be assumed. We assumed a constant symmetric variance (σ_G^2) for both peaks of the overlapped pair; this is consistent with the scenarios described above. Thus for Gaussian peaks the total variance (σ_G^2) was fixed, whereas for the EMG peaks the total variance ($\sigma_G^2 + \tau^2$) increased as the asymmetry ratio (τ/σ_G) was increased.

The generation of overlapping peaks is illustrated in Fig. 1. Isolated symmetric



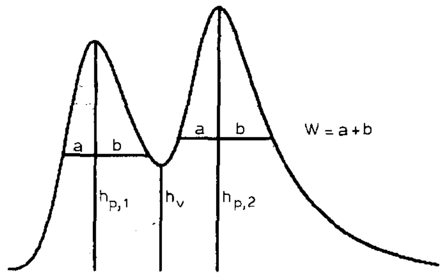


Fig. 1. Overlapping peak simulations. Individual Gaussian and exponentially modified Gaussian (EMG) functions in Fig. 1a were offset and added to produce the symmetric peak–symmetric peak, tailed peak–tailed peak combinations shown in Fig. 1b, as well as the tailed peak–symmetric peak and tailed peak–symmetric peak combinations that are not shown.

Fig. 2. Graphical parameters of overlapping peaks, shown here for two tailed (EMG) peaks of equal area with an asymmetry ratio (τ/σ_G) of 2. The relative valley is defined as the ratio of h_v/h_p, and is conveniently expressed as a percentage. The width, W, and asymmetry factor, b/a, are shown at the peak height fraction, $\alpha = 0.5$.

(Gaussian) and tailed (EMG, $\tau/\sigma_G = 2$) peaks of unit area in Fig. 1a are merged with identical peaks to produce the symmetric–symmetric and tailed–tailed overlapping peak combinations in Fig. 1b. The resolution, R_s, defined as $\Delta t_R/4\sigma^2$ [where σ^2 is the total variance (second statistical moment)], was 0.625 for both pairs of overlapping peaks.

Measures of peak overlap. Fig. 1b illustrates the inadequacy of the present definition of resolution to describe the overlap of real chromatographic peaks. As noted long ago by Kirkland *et al.*[17], the "apparent resolution" is better for the tailed peaks than for the symmetric peaks. This paradox can be explained in terms of the greater error in peak height and peak area estimation for the tailed peak pair which we discuss in detail shortly.

Given the inadequacy of the resolution parameter to describe peak overlap, we decided to employ an empirical parameter, the relative valley, as illustrated in Fig. 2. We define the relative valley as the ratio of the height of the valley, h_v, to the apparent height of the peak in question, h_p. We will frequently report it as a percentage, *i.e.*, % (relative) valley = $h_v/h_p \times 100$. Note that, unless the apparent peak heights are the same for two peaks of an overlapping peak pair ($h_{p,1} = h_{p,2}$), the relative valleys for the two peaks will not be the same. Finally, although the relative valley is empirical, it is an unambiguous parameter, one that is easily measured in practice.

Errors in peak area and peak height measurements

General comments. The results of our study show that errors in the perpendicular-drop and apparent peak height methods for the measurement of peak area and peak height of overlapping peaks are due primarily to two distinguishable effects, one resulting from a difference in the relative size (areas) of the two peaks that overlap, the other resulting from the asymmetry of one or both peaks of the overlapped pair. We shall denote these effects as *size effects* and *asymmetry effects*, respectively. Whereas the error in peak area will be negative for one peak and positive for the other, the error in peak height will be positive for both peaks. Obviously the error in either parameter becomes larger as the degree of peak overlap (relative valley) increases.

Size effects. We examined the errors in peak height and peak area measurement as a function of the relative areas of overlapping symmetric (Gaussian) peaks (and the degree of peak overlap). Because size effects have already been thoroughly discussed[9,10], we only summarize the results for purposes of comparison with the asymmetry effects.

For all degrees of peak overlap (relative valley), errors in peak height are much smaller than corresponding errors in peak area. Errors in peak height are negligible, in fact, whenever the valley is less than 50%. For a given degree of peak overlap, errors in peak area are much more dependent on the area ratio than errors in peak height. The relative error in peak area and peak height was larger for the smaller of the overlapping peaks. Finally, the maximum errors due to size effects for peaks with an area ratio of four and a perceptible valley were −28% for peak area and +7.5% for peak height (for the smaller peak in both cases).

Asymmetry effects. We examined the errors in peak height and peak area measurement as a function of the asymmetry (tailing) of one or both of the overlapping peaks, *i.e.*, by examining tailed peak–tailed peak, tailed peak–symmetric peak, and symmetric peak–tailed peak combinations for peaks of equal size. Except for a recent reference to peak height[18], asymmetry effects have been completely ignored.

Since we assumed a constant symmetric variance (σ_G^2) for all peaks, tailed peaks of unit area are shorter than corresponding symmetric peaks, as shown in Fig. 1a. It was thus impossible to produce overlapping peaks of equal area *and* equal peak height for the tailed peak–symmetric peak and symmetric peak–tailed peak combinations. Because area is more easily fixed for EMG peaks (see eqn. 2), our studies used peaks of equal area.

(1) Tailed peak–tailed peak combinations. An example of the asymmetry effect is given in Fig. 3 for the tailed peak–tailed peak peak combination with $\tau/\sigma_G = 3$. Individual peaks were lightly traced, whereas the composite peak is shown with a

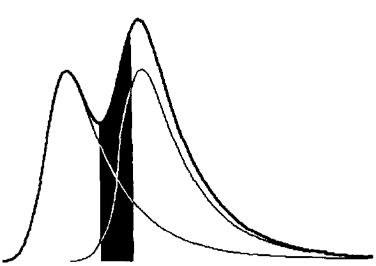


Fig. 3. Illustration of the asymmetry effect for the tailed peak/tailed peak combination, shown here for two overlapped EMG peaks of equal area with an asymmetry ratio (τ/σ_G) of 3. Shaded area represents the error of the perpendicular-drop algorithm for peak area. The errors in peak area are −34% and +34% for the first and second peaks; the errors in peak height are 0.03% and 26%.

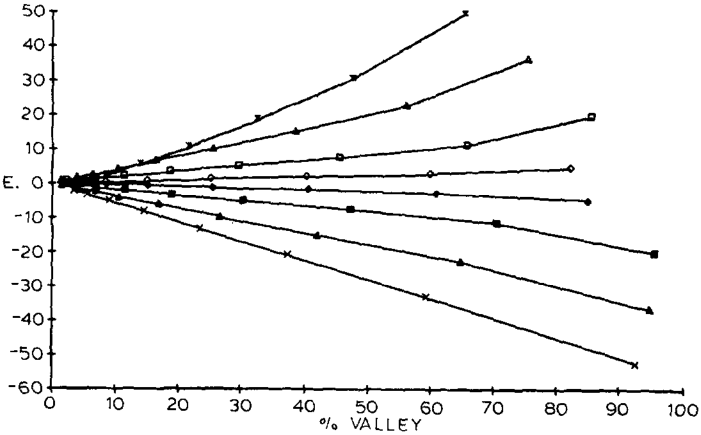


Fig. 4. Summary of asymmetry effects on peak area for two overlapping, tailed peaks of equal area. Filled symbols indicate results for the first peak, open symbols represent the second peak. Values for τ/σ_G (asymmetry ratio): 0.5 (○), 1 (□), 2 (△), and 4 (×).

bold tracing. The shaded area represents the error of the perpendicular-drop algorithm, *i.e.*, the amount by which the area of the first peak is underestimated and by which the area of the second peak is overestimated. The shaded area represents a substantial portion (34%) of the total area. Also, in contrast to the size effect, the asymmetry effect on peak *height* is significant (+26%), although only for the second peak.

Asymmetry effects for the tailed peak–tailed peak combinations are summarized in Figs. 4 and 5 for peak area and peak height for τ/σ_G = 0.5, 1, 2, and 4. As expected, the greatest errors in peak area occurred for the peak combinations with τ/σ_G = 4 (the greatest tailing). The maximum observed errors were −52% and +50% for the first and second peaks. (Slight differences are attributed to truncation errors in the integration of the second peak.) For peak height, the greatest errors for

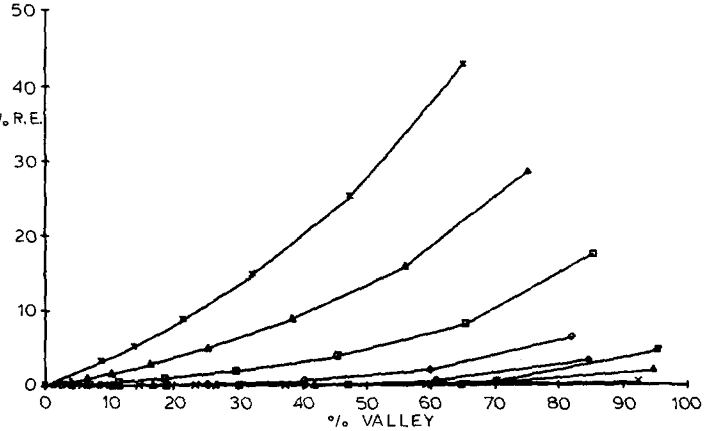


Fig. 5. Summary of asymmetry effects on peak height for two overlapping, tailed peaks of equal area. Conditions as in Fig. 4.

the first peak (+5%) occurred for peak combinations with moderate tailing ($\tau/\sigma_G = 1$) whereas the greatest errors for the second peak +40%) occurred for the combination with maximum tailing.

(2) Tailed peak–symmetric peak combinations. The asymmetry effects for this combination were slightly larger for peak area and moderately lower for peak height than for the tailed peak–tailed peak combination above, although the general pattern remained the same. For this reason no figures are included for these data, although they will be made available as supplementary material. The maximum errors in peak area were −58% and +56% for the first and second peaks. For peak height, the largest errors were +5% for the first peak (at $\tau/\sigma_G = 0.5$) and +20% for the second peak (at $\tau/\sigma_G = 4$). The difference between the asymmetry effects for the tailed peak–tailed peak and tailed peak–symmetric peak combinations is attributed to the fact that, although the first and second peaks were of equal area for both combinations, the first peak was shorter in the tailed peak–symmetric peak combination.

(3) Symmetric peak–tailed peak combinations. The results for the case where only the second peak is tailed were different from the results of the two previous combinations. Errors in peak area were much lower (±12%), and the area of the first peak was *over*estimated instead of *under*estimated as in the previous combinations (and vice versa for the second peak). Errors in peak height were comparable for the first and second peaks (a maximum of +8%), and fell between the errors observed for the first and second peaks in previous combinations. Finally, errors in both peak area and peak height were much less dependent on the asymmetry of the second peak.

The difference between the asymmetry effects for the symmetric peak–tailed peak combination and the first two combinations was not unexpected. By definition the leading edge of a tailed (EMG) peak more closely resembles a Gausian peak than the trailing edge. Since the leading edge of the EMG peak is the part that overlaps with the Gaussian in the symmetric peak–tailed peak combination, we expect the asymmetry effects to be much less pronounced. The effects we do observe are attributed to the somewhat smaller slope of the leading edge of the EMG peak due to the larger variance. We could predict similar results for two partially resolved Gaussian peaks, the second of which had a larger variance.

In summary, the most serious errors in peak area and peak height resulting from asymmetry effects occurred when the first peak was tailed. These errors, which have previously been ignored, sometimes exceeded 50%. When only the second peak was tailed, the errors were less than 15%.

Combined size–asymmetry effects

The combined effects of size and asymmetry were studied by examining the overlapping peak combinations tailed peak–tailed peak, tailed peak–symmetric peak, and symmetric peak–tailed peak for peaks with area ratios of 0.25:1, 0.5:1, 1:0.5, and 1:0.25. For reasons of brevity, we will not consider all combinations in detail. Data and charts are available upon request.

For peak area, in general, the errors induced by the combined size–asymmetry effects were additive, and thus substantially larger than for the separate effects, although occasionally a fortuitous cancellation of error was observed. (A cancellation of error is possible, since both effects produced positive and negative errors.) For

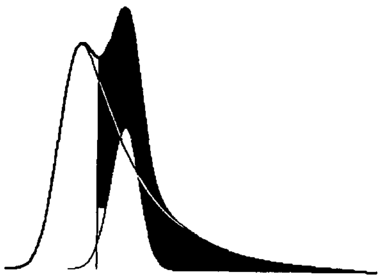


Fig. 6. Illustration of the combined effects of size and asymmetry on the measurement of peak area and peak height. Relative peak area (tailed peak: symmetric peak) = 4:1. Shaded area represents the error of the perpendicular-drop algorithm for peak area. The errors in peak area are −54% and +208% for the tailed and symmetric peaks respectively; errors in peak height are +0.5% and +86.0, respectively.

peak height, since errors due to size and asymmetry are always positive, no cancellation of error is possible. In general, errors in peak height due to combined effects were additive, but since size effects are several times smaller than asymmetry effects, the errors in peak height due to combined effects were only marginally larger than errors due to asymmetry effects alone.

The largest errors observed for both parameters resulted from the overlap of a large highly tailed peak with a small, symmetric peak, as shown in Fig. 6 for peaks with an area ratio of four-to-one. As in Fig. 3, the shaded area represents the error of the perpendicular-drop algorithm, *i.e.*, the amount by which the area of the tailed peak is underestimated and by which the area of the symmetric peak is overestimated. For the tailed and symmetric peaks in Fig. 6 the errors in peak area were −54% and +208%; errors in peak height were +0.5% and +86%.

Improving the accuracy of quantitation

The above discussion shows that, due to symmetry effects which have previously been ignored, the systematic errors of current integrator-based methods for peak area and peak height measurement on overlapping peaks are much larger than previously estimated. This is particularly true for partially resolved peaks of nearly equal area, for which errors in the perpendicular-drop algorithm have formerly been assumed to be negligible.

The best way to eliminate quantitative errors caused by overlapping peaks is to eliminate the overlap, *i.e.*, change the chromatographic conditions so that the peaks of interest are baseline-resolved. Because this cannot always be done in practice, it is thus desirable to consider what improvements in quantitation are possible wihtout an improvement in the separation.

Empirical area equation. Since the errors in peak height and peak area measurement for overlapping peaks are a direct result of the distortion caused by the overlap, a logical approach would be to use or develop a quantitative method, based on measurements in regions of minimum distortion. In our studies, we observed much less distortion for the first peak than the second, as evidenced by the generally insignificant errors in peak height for the first peak.

We recently developed[19] a set of empirical equations for the calculation of

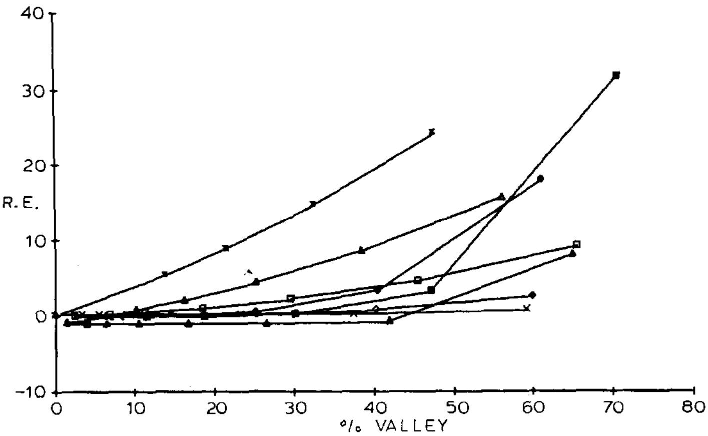


Fig. 7. Summary of asymmetry effects on eqn. 3, $A = 1.64\, h_p\, W_{0.75}\, (b/a)^{+0.717}$, for two overlapping tailed peaks of equal area. Conditions as in Fig. 4.

peak area for symmetric (Gaussian) and tailed (EMG) peaks. These equations were obtained by fitting plots of A_{true}/A_G *vs.* the asymmetry factor (b/a), where A_G is the area calculated using a Gaussian equation. Although these equations were originally developed for peak modeling studies, one of them proved useful for the present study. The equation is

$$A = 1.64\, h_p\, W_{0.75}\, (b/a)^{0.717} \quad (3)$$

where A is the peak area, h_p is the peak height, and $W_{0.75}$ and b/a are the peak width and asymmetry measured at 75% of the peak height, respectively. The bias of eqn. 3 is less than ±1% for well-resolved symmetric (Gaussian) and tailed (EMG) peaks from $\tau/\sigma_G = 0$ to 4.2. The expected precision of eqn. 3, calculated via error propagation, is ±1%, assuming a precision of ±0.5%, ±0.5%, and ±1% for h_p, $W_{0.75}$, and b/a.

The accuracy of eqn. 3 for overlapping peaks is shown in Fig. 7 for the tailed peak–tailed peak peak combination. For peak overlap (valleys) up to 45%, errors were less than ±4% for both the first and the second peaks, except in the cases of extreme tailing ($\tau/\sigma_G = 3, 4$). In these instances, the relative errors were large for the second peak but less than ±1% for the first peak.

Eqn. 3 was comparably accurate for the remaining tailed peak–tailed peak combinations and for all the tailed peak–symmetric peak combinations. Errors for the first peak never exceeded ±4%; errors for the second peak were large, as in Fig. 7, for cases of extreme tailing ($\tau/\sigma_G = 3, 4$).

For the symmetric peak–tailed peak combinations, eqn. 3 was more accurate for the second peak than for the first. In addition, the errors were nearly independent of the degree of tailing. For valleys up to 50%, errors were less than ±2% for the second (tailed) peak regardless of asymmetry. For the first (symmetric) peak, errors ranged from +5 to +20%.

Although we did not examine peak combinations with area ratios outside the range of 1:4 to 4:1, we can safely assume that, for the *larger* peak of an overlapped pair, eqn. 3 wil be more accurate than stated above (provided the valley limits are not exceeded), because the distortion of the larger peak of an overlapped pair for peak combinations outside the range of 1:4 to 4:1 will be less than for combinations within this range.

To summarize, for overlapping peaks with area ratios between 1:4 and 4:1, the empirical area equation, eqn. 3, is accurate to within $\pm 4\%$ for the first peak of tailed peak–tailed peak or tailed peak–symmetric peak combinations, provided that the valley between peaks is less than 45%. For the symmetric peak–tailed peak combination, eqn. 3 is accurate to within $\pm 2\%$ for the second peak, if the valley is less than 50%. For overlapping peaks with area ratios outside the range of 1:4 and 4:1, eqn. 3 will be somewhat more accurate, but only for the larger peak of the overlapped pair.

Combined area equation–integrator approach. By itself, eqn. 3 facilitates the accurate ($< \pm 4\%$) quantitation of only *one* peak of an overlapped pair. However, if an integrator is used to measure the total area of the two overlapping peaks, the area of the other peak can be accurately determined by subtraction, *i.e.*,

$$A_{\text{other}} = A_{\text{T}} - A_{\text{eq}} \tag{4}$$

where A_{T} is the total area provided accurately by the integrator, A_{eq} is the area of the peak for which eqn. 3 is accurate, and A_{other} is the area of the other peak.

The accuracy and precision by which A_{other} can be measured is, of course, related to the accuracy and precision by which A_{T} and A_{eq} can be measured. Assuming (1) A_{T} can be measured with much greater accuracy and precision than A_{eq}; and (2) neglibible covariances, the relative error (R.E.) and relative standard deviation (R.S.D.) of A_{other} are given by

$$\text{R.E.}\ (A_{\text{other}}) = -\ [A_{\text{ratio}}\ \text{R.E.}\ (A_{\text{eq}})] \tag{5}$$

$$\text{R.S.D.}\ (A_{\text{other}}) = A_{\text{ratio}}\ \text{R.S.D.}\ (A_{\text{eq}}) \tag{6}$$

where $A_{\text{ratio}} = A_{\text{eq}}/A_{\text{other}}$. Thus, the accuracy and precision of the indirectly measured

TABLE I

RELATIVE ERROR IN PEAK HEIGHT AND PEAK AREA MEASUREMENT FOR A LARGE, TAILED PEAK, OVERLAPPED WITH A SMALL, SYMMETRIC PEAK*

Parameter	*Relative error (%)*	
	Tailed peak	*Symmetric peak*
Peak height	0.0	+ 45.6
Peak area (perpendicular-drop)	−31.1	+117.3
Peak area, eqns. 3 and 4	+0.4	− 1.6

* Tailed (EMG) peak with $\tau/\sigma_{\text{G}} = 4$ and area = 1. Symmetric (Gaussian) peak with area = 0.25. Valley between the peaks is 54% relative to the tailed peak (60% relative to the symmetric peak).

peak depends directly on the accuracy and precision of the area equation and on the area ratio of the overlapping peaks. For $0.25 < A_{ratio} < 1$, the relative accuracy and precision of A_{other} will be better than for A_{eq} (the accuracy of A_{eq} was not confirmed for $A_{ratio} < 0.25$). For $A_{ratio} > 1$, the accuracy and precision of A_{other} will be worse. Assuming R.E. (A_{eq}) = R.S.D. (A_{eq}) = $\pm 1\%$, A_{ratio} can be no higher than 10:1 for acceptable ($\pm 10\%$) accuracy and precision of A_{other}.

As described above, the combined equation–integration method for peak area measurement of overlapping peaks (eqns. 3 and 4) is appropriate for $0.25 < A_{ratio} < 10$. Although limited to pairs of peaks with area ratios in this range, this method can easily be implemented on programmable integrators and data systems and will improve the accuracy of area quantitation substantially in these cases. An example of the superior accuracy of the combined area equation–integration method for peak quantitation (eqns. 3 and 4) over common integrator–data system methods for peak height and peak area is shown in Table I for a highly tailed peak ($\tau/\sigma_G = 4$, area = 1) overlapped with a symmetric peak (area = 0.25). The peak combination is the same as illustrated in Fig. 6, exept that the peak overlap was somewhat less (valleys of 54% and 60% for the tailed and symmetric peaks).

Correction factors. The true value, T, for peak area or peak height, is related to the observed value, O, by the relative error, R.E. = $(O-T)/T$. Solving this equation in terms of T yields

$$T = O/(\text{R.E.} + 1) \tag{7}$$

We can regard $1/(\text{R.E.} + 1)$ in eqn. 7 as a correction factor (C.F.) and write

$$T = O \times \text{C.F.} \tag{8}$$

Correction factors can be interpolated manually or by computer from a previously generated look up table. Alternatively, if a sufficiently accurate functional relationship can be obtained, these correction factors can be calculated on-line using a computer.

When *only* the size effect is considered (Gaussian peak shapes are assumed), it is possible to calculate accurate correction factors as a function of only two variables: area ratio and degree of peak overlap (resolution, R_s), *i.e.*,

$$\text{C.F.} = \text{f}[A_{ratio}, \text{peak overlap } (R_s)] \tag{9}$$

Correction factors based only on size effects have been calculated previously for peak areas[11,12]. However, as we have shown in this report, asymmetry effects on the quantitation of overlapping peaks are frequently as important or more important than size effects. Thus, size-effect-only "correction factors" are generally inaccurate.

Unfortunately, the incorporation of peak asymmetry into the correction factor, as shown in eqn. 10, makes the calculation and use of correction factors difficult if not impossible.

$$\text{CF} = \text{f}[A_{ratio}, \text{type of combination, peak overlap } (\neq R_s), \text{peak asymmetry}] \tag{10}$$

First, for a given set of area ratios the number of overlapping peak combinations has increased from one (symmetric–symmetric) to four (symmetric–symmetric, tailed–tailed, tailed–symmetric, and symmetric–tailed). Since each combination requires a pair of correction factors (one for each peak of the overlapped pair), eight sets of correction factors are now required instead of two. Second, an empirical measure of peak overlap must be used, since the fundamental resolution parameter, R_s, can no longer be visually estimated (except by comparison with standard drawings for the hundreds of possible combinations, which is clearly impractical). Finally, the dimensionality of each set of correction factors is increased from two to four. If correction factors are to be *interpolated* from tabulated values, a four-dimensional table will be required instead of a two-dimensional table for each set of correction factors. If correction factors are to be *calculated,* an accurate relationship for eqn. 10 must be deduced, presumably by multivariate regression. (Our efforts in obtaining a satisfactory empirical relationship for eqn. 10 were unsuccessful.)

Other discussion

Implications for preparative liquid chromatography. Because asymmetry effects have previously been ignored, peak purity has probably been overestimated for overlapping tailed peaks. In particular, if the valley is used as the cutpoint for peak fractions of overlapping peaks, the second compound will be highly contaminated with the first. To ensure the purity of the second peak, the cutpoint should be made more conservatively, *i.e.,* considerably later than the valley between the peaks.

Errors due to overlap of more than two peaks. Although we did not perform a simulation involving a *series* of overlapping peaks, the trends observed for overlapping *pairs* of peaks enables us to make the following predictions for a series of peaks, assuming that (1) the tailing is approximately the same for all peaks; (2) the peak areas are within a factor of 4 for neighboring peaks; and (3) all neighboring peaks have valleys between them.

For the first peak of a series, the error in peak height wil be negligible, as will the error in peak area calculated using eqn. 3, if the valley is less than 45%. In contrast, the error in peak area, as determined by the perpendicular-drop algorithm, wil be large, especially if the peaks are highly tailed. For interior peaks, errors in peak height are expected to be significant, due to peak distortion from both sides. Errors in peak area resulting from use of eqn. 3 will be also be substantial. On the other hand, errors in the perpendicular-drop algorithm for peak area may or may not be significant. Asymmetry effects are expected to cancel, whereas size effects may be somewhat enhanced if the middle peak is smaller than its neighbors. Finally, for the last peak in a series of partially resolved tailed peaks, we expect all three measurements (peak height, peak area via perpendicular-drop, and peak area via eqn. 3) to be inaccurate. Note that an approach analogous to eqn. 4 (eqn. 11 below) will not work, because A_{interior}, the area of the interior peaks, cannot always be determined with sufficient accuracy.

$$A_{\text{last}} = A_{\text{T}} - A_{\text{first}} - A_{\text{interior}} \tag{11}$$

CONCLUSIONS

Because peak asymmetry (tailing) has previously been ignored, the accuracy of the integrator-based measurement of peak area and peak height has been substantially overestimated for overlapping peaks. For peaks of comparable size, errors in the perpendicular-drop algorithm for peak area, which were formerly assumed to be negligible, can be greater than 50%. Errors in peak height, also frequently assumed to be negligible, can be greater than 40%. The combination of asymmetry effects and differences in relative peak size can lead to errors in peak height that exceed 80% and errors in peak area that exceed 200%. Baseline or near-baseline resolution is therefore essential for the accurate quantitation of overlapping peaks when common integrator–data system methods are used and one or both peaks is tailed. Alternatively, the combined use of an empirical area equation and normal integration procedures (eqns. 3 and 4) allows accurate quantitation ($< \pm 4\%$) for overlapping peaks, provided that the valley between the peaks is less than 45%. The use of correction factors does not appear to be a viable approach whenever one or both overlapping peaks is tailed.

SUPPLEMENTARY MATERIAL

Data and charts for all the overlapping peak combinations described are available upon request.

REFERENCES

1 J. M. Davis and J. C. Giddings, *Anal. Chem.*, 57 (1985) 2168–2177.
2 J. M. Davis and J. C. Giddings, *Anal. Chem.*, 57 (1985) 2178–2182.
3 R. F. Lacey, *Anal. Chem.*, 58 (1986) 1404–1410.
4 S. D. Frans, M. L. McConnel and J. M. Harris, *Anal. Chem.*, 57 (1985) 1552–1559.
5 R. A. Vaidya and R. D. Hester, *J. Chromatogr.*, 287 (1984) 231–244.
6 R. A. Vaidya and R. D. Hester, *J. Chromatogr.*, 333 (1985) 152–155.
7 A. W. Westerberg, *Anal. Chem.*, 41 (1969) 1770–1777.
8 J. A. Perry, *Introduction to Analytical Gas Chromatography*, Marcel Dekker, New York, 1981, pp. 361–385.
9 L. R. Snyder and J. J. Kirkland, *Introduction to Modern Liquid Chromatography*, Wiley, New York, 2nd ed., 1979, pp. 38–48, 556–559.
10 L. R. Snyder, *J. Chromatogr. Sci.*, 10 (1972) 200–212.
11 J. Novák, K. Petrović and S. Wičar, *J. Chromatogr.*, 55 (1971) 221–229.
12 E. Proksch, H. Bruneder and V. Granzner, *J. Chromatogr. Sci.*, 7 (1969) 473–483.
13 D. Hanggi and P. W. Carr, *Anal. Chem.*, 57 (1985) 2394–2395.
14 R. Delley, *Chromatographia*, 18 (1984) 374–382.
15 R. Delley, *Anal. Chem.*, 57 (1985) 388.
16 J. P. Foley and J. G. Dorsey, *J. Chromatogr. Sci.*, 22 (1984) 40–46.
17 J. J. Kirkland, W. W. Yau, H. J. Stoklosa and C. H. Dilks, Jr., *J. Chromatogr. Sci.*, 15 (1977) 303–316.
18 D. A. Dezaro, T. R. Floyd, T. V. Raglione and R. A. Hartwick, *Chromatogr. Forum*, 1 (1986) 34–37.
19 J. P. Foley, *Anal. Chem.*, submitted for publication.