ELSEVIER

Journal of Chromatography A, 842 (1999) 321–340

JOURNAL OF CHROMATOGRAPHY A

Review

Peak distortion, data sampling errors and the integrator in the measurement of very narrow chromatographic peaks

Norman Dyson

Dyson Instruments, The Croft, Kellfield Avenue, Low Fell, Tyne and Wear NE9 5YP, UK

Abstract

As more chromatographers consider using the techniques of fast chromatography or capillary electrochromatography, the instrumentation problems of generating and measuring these very narrow peaks is re-appraised. In particular the general view that 20–30 samples per peak is sufficient data to measure a peak is shown to be true under limited circumstances. A reworking of numerical integration theory and calculation of the errors of Newton Cotes rules when applied to chromatographic peaks, shows that asymmetry creates more subtle mischief: as many as 350 samples/peak may be needed to achieve 0.1% accuracy. Specialist, low time constant units are required to generate narrow peaks and a new breed of fast sampling data processor is required to measure them. As peaks narrow, it is increasingly important that data processors help analysts to identify data under-sampling by reporting peak asymmetry, actual sampling frequency (as opposed to that initially programmed) and number of measured samples/peak as a combined validation diagnostic. Finally the article considers the lack of development of new deconvolution based data processors and points towards the lack of information inside flame ionisation and ultraviolet absorbance detectors. If new data processors were to become available for information rich detectors, their benefits: improved accuracy, precision and greater confidence in results, would have to be weighed against the costs of adopting them, re-working of analysis methods and retraining staff. Many production laboratories would find it uneconomical and would stay with the old methods.

Keywords: Peak shape; Data sampling error; Fast chromatography; Electrochromatography; Data processing

Contents

E-mail address: ndyson@aol.com (N. Dyson)

0021-9673/99/$ – see front matter
PII: S0021-9673(99)00299-X

1. Introduction

Peaks generated by fast chromatography (FC) or capillary electrochromatography (CEC) are no different to measure than other chromatographic peaks, just narrower. But narrow FC peaks cannot be generated, for example, by installing a short column into a standard capillary GC, the injection port and electronic time constants will frustrate the work of the column. Specialist units are required which contribute little variance to observed peak width. In any case, to be measured accurately, very narrow peaks need a higher sampling frequency than current laboratory data processors can offer. For very narrow peaks a new generation of data processors is needed which can sample data at frequencies up to 6 kHz. This will have a knock on effect on data storage and noise suppression.

The number of samples/peak to achieve a required measurement accuracy (in any type of chromatographic peak, broad or narrow) is reconsidered here. It tends to be assumed that existing integrators can sample the detector signal fast enough and it only needs about 20 to 30 samples/peak for accurate peak measurement, but neither assumption is necessarily true. It is shown below that an asymmetric peak might require as many as 350 data samples to be measured at 0.1% accuracy. If the data processor has a typical maximum sampling frequency of 20 Hz, peaks narrower than 17 s base width cannot be measured at this accuracy. On the other hand, the 20 Hz sampling frequency time averages and suppresses typical background noise and if data sampling frequencies are going to increase by an order of magnitude or more, alternative noise suppression techniques will have to be used. The basic evaluation of measurement errors using Newton Cotes rules of peak measurements uncovers yet another disadvantage of peak asymmetry: asymmetric peaks need more samples/peak than symmetrical ones to achieve the same measurement accuracy.

Finally, the shortcomings of integrator algorithms are reviewed and the reasons why data processors using better algorithms have not emerged are considered. If an integrator using accurate deconvolution techniques became available, the task and cost of re-training staff and re-validating experiments would be considerable [1]. This new integrator would need to be coupled to a ''three dimensional detector'' such as a GC–MS or diode array, because flame ionisa-

tion and ultraviolet absorbance detectors do not carry enough information to allow true deconvolution. The adoption costs might be so high and the benefits so intangible that labs simply stay with their old methods arguing that they are sufficient.

2. Peak generation and shape distortion

Not all analyses require long columns or hundreds of thousands of theoretical plates. Some complex samples do, but other samples can be analysed on shorter columns which are cheaper to buy (sometimes) and provide shorter analysis times, which improves productivity. CEC columns are so efficient they provide large numbers of theoretical plates and short analysis times. Typical capillary GCs have been designed to use capillary columns of not less than about 10 to 15 m. The shortest usable column depends on the injector, detector amplifier time constant and therefore on the age of the GC to some extent. Existing ''off the shelf'' data processors have a maximum sampling rate of about 20 samples/s maximum which is fast enough to measure symmetrical peaks of 2 to 3 s base width but not much else. Installing a short column (less than 5 m, say) into a typical bench GC system will produce shorter analysis times but it will not achieve the peak resolution that the column specification might suggest is possible because off-column broadening will undo the work of the column.

In 1966 Sternberg [2] described how column performance is degraded by the design of the injection port, dead volume and surface activity in the flow lines, detector cell design and electronic time constants. All of these factors contribute to observed peak width, or more accurately to peak variance, $\sigma^2_{\text{observed}}$. We know too that the processes of data sampling, integration and signal display also make a contribution to observed variance [3]. In rigorous terms, σ is the standard deviation of a symmetrical peak so that σ^2 is the variance of a symmetrical peak, but it is also common to use σ^2, suitably labelled, to be the variance of an asymmetric peak and equal to the second moment – the true variance of an asymmetric peak, and that is how the term is used immediately below. Later in this review, the variance of an exponentially modified gaussian peak is given by $(\sigma^2+\tau^2)$ after Foley and Dorsey.

Sternberg showed the variance or second moment of a peak to be the sum of component variances:

$$\sigma^2_{\text{observed}} = \sigma^2_{\text{column}} + \sigma^2_{\text{extra-column}} \quad (1)$$

$$= \sigma_{\text{column}} + \sigma^2_{\text{injector}} + \sigma^2_{\text{detector}} + \sigma^2_{\text{pipework}} \quad (2)$$

The second moment of a peak, is measured by time sampling the detector signal as the peak emerges from the column. It can be measured graphically from peak width by assuming a peak shape such as gaussian where peak width at half height, $w_{1/2}$, is equal to 2.355σ, and then variance is equal to $(w_{1/2})^2/5.546$.

The importance of Eq. (2) is that component variances not only add linearly, but they can be considered separately and reduced or removed by improved design. Eq. (2) shows that total peak width builds up like onion rings from component contributions, although all the analyst ever sees is the outer layer, the peak envelope.

In an ideal world σ_{column} is small and yet still the dominant term on the right hand side of Eq. (2). As column length and phase loading decrease, σ^2_{column} becomes smaller and other terms, especially $\sigma^2_{\text{injector}}$ begin to dominate. There comes a point when the peak width is determined more by chromatograph design than by the column.

The contribution to peak variance by the pipework is eliminated by removing any dead volume where analyte molecules might become trapped or emerge slowly, and by ensuring that off column tubing does not interact with the sample.

The contribution of the detector is minimised by making the detector cell geometrically small and streamlined, and by designing the time constant of the electronics to be as small as possible consistent with its other role as a noise filter. In a flame ionisation detection (FID) system where the tip of a capillary column sits just below the flame, pipework is eliminated and the cell volume is effectively zero.

The biggest contribution to peak variance is made by the injection port where sample volatilises, mixes with carrier gas and is delivered to the column as a vapour plug spread over several centimetres of column. This broadening will dissipate and may even disappear if the column is long enough. The column refocuses the analyte zones into narrower bands by slowing down the leading edge of each band, by

adsorption, while the trailing edge continues to move forward and catch up. There is a rolling up effect, a narrowing of the band until it equilibrates with the column and adopts the geometry it would otherwise have had without injector dispersion. But this takes time, and if the analyte elutes before equilibration is complete the peak will retain some of the original injector band broadening. See Fig. 1 which shows the elution of 1 ml of valve injected gaseous hydrocarbons from a KCl/alumina capillary column after injection from a poor sampling valve.

In Fig. 1 the peak widths decrease with increasing retention which is the opposite of expectation, it suggests that column efficiency is increasing with time. What is really happening is that dispersion introduced by the injection port is dissipating with time. Short columns are clearly more likely to show this effect than long ones.

Golay [4] worked out the equation of column efficiency in terms of gas phase flow rate and column processes. This equation was extended by Gaspar et al. [5] to include extra column broadening. For a capillary column, the height, H, of a theoretical plate is related to the linear velocity, u, of carrier gas through the column by:

$$H = \frac{B}{u} + (C_l + C_g)u + Du^2 \tag{3}$$

where B/u is the longitudinal dispersion term, $(C_l + C_g)u$ is the column mass transfer term, and D is the extra column contribution term:

$$D = \frac{\tau^2}{(1+k)^2} \times L \tag{4}$$

where τ = the combined system time constant from

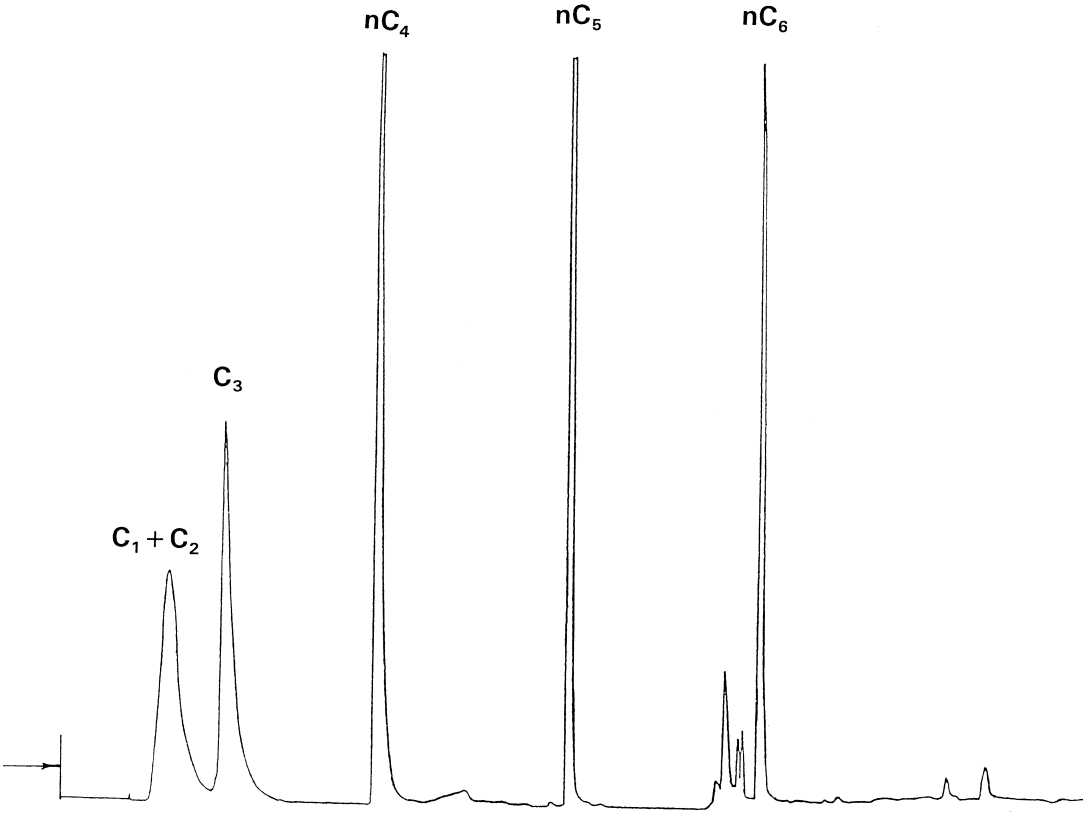


Fig. 1. Injection port dispersion will dissipate if retention is long enough.

all extra column broadening, k = partition ratio of measured analyte, and L = column length.

The D term in Eq. (4), the *extra*-column term it will be noted, is inversely dependent on column length, L, and solute retention, $(1+k)^2$, both of which are column parameters. The shorter the column and the less the retention of solutes, the greater will be the value of D. This makes sense if the column is considered to be 'repairing' the peak shape distorted by the injection port. Clearly, post column distortion cannot be so repaired and remains in the final peak shape. When short columns are used, a large part of residual extra-column broadening is injector broadening which has not had time to dissipate. This leaves the analyst with the choice either to accept this broadening or to restrict sample volume to minimise the injection pulse, and this in turn limits the minimum detectable quantity and usable linear dynamic range.

There is a curious intermediate stage where a short length of high-performance capillary column should deliver peaks which are too narrow for a data processor to measure accurately at 20 Hz sampling frequency, but the extra-column broadening expands these peaks into the width range where the data processor can sample them sufficiently and measure them accurately. This happens when there are few sample components and the column, however short, has surplus theoretical plates for the required separation. In other words some of the baseline between peaks is sacrificed to peak broadening. There need be no loss of analytical information if only sample composition is required.

τ is a single term representing all of the extra column broadening mechanisms. As Eqs. (1) and (2) suggest, it can be related to its components by:

$$\sigma^2_{\text{extra-column}} = \sum \lambda_{\text{ec}}.\tau^2_{\text{ec}} \qquad (5)$$

where τ_{ec} is the time constant of each extra-column broadening mechanism and λ_{ec} are individual proportionality constants for each mechanism.

In contrast to Eq. (5), the exponentially modified gaussian (EMG) function has only one time constant to represent all mechanisms. Torres-Lapasio et al. [6] attempted to represent the separate mechanisms by defining a polynomially modified gaussian (PMG) and had some success though they were limited to three factors and these factors were not independent of each other.

2.1. The size of fast peaks: how small is "narrow"?

Annino [7] minimised the extra column broadening described by Steinberg. He compared the extra-column time constants of manual and auto-injection and found that changing to auto-injection (of 1 μl sample volume) reduced the time constant from 700 ms (manual) to 45 ms. These figures include any time constant contributions from the detector, but they clearly demonstrate that auto injection is the better technique. What was better still was the thermal pulse cryo-injector of Klemp et al. [8]; designed specifically for fast injection, it has a system time constant measured at 10 ms. Annino used a 3.4 m×0.25 mm DB-1 column, a detector amplifier with a 5 ms time constant and the cryo-injector, and generated peaks whose variance was degraded (increased) by no more than 2% (Steinberg's ϕ term). He compared three commercial GC systems and at best, the resulting peaks were as narrow as 0.19 s base width. On a different GC system Annino [9] produced a methane peak of about 0.13 s base width.

A brief literature survey determined that Beens et al. [10] used 2 m and 0.5 m×0.10 mm I.D. columns in a two-dimensional GC system to create peaks of <0.5 s base width. Kutter et al. [11] used a solvent programmed microchip open channel electrochromatograph to separate four major and four minor peaks in 5 s, again the base width is about 0.5 s.

Dadoo et al. [12] also used CEC to separate five polycyclic aromatic hydrocarbons in about 1 s. The narrowest peak was quoted to have a half width of 80 ms, and assuming a gaussian shape, this equates to 200 ms, or 0.2 s, base width. The time constant of their detector amplifier was set at 3 ms and the data sampling period was 8.3 ms. (sampling frequency of 120 samples/s).

Today's technology therefore seems able to produce peaks of about 0.1 to 0.5 s base width. If a designer were to target a peak width and design a data processor to measure it, that peak would be no broader than 0.1 s. To give the product some life time the designer would also consider what is likely

to be measured in 5 to 10 years time. Allowing for future developments, especially in the fields of CEC and 'microchip instruments', 0.01 s base width does not seem an unreasonable target. This base width is firmly in the 'noise domain' of current integrators and such peaks would not be measured, peak recognition algorithms which separate the broad from the narrow would filter them.

3. The technology of sampling: A/D conversion

In 1973 Kelly and Horlick [13] summarised the rules for using analog to digital (A/D) converters and the problems likely to arise when an A/D converter is used at the edge of its performance. In spite of its age, this paper is still relevant.

There is a finite time between the moment analyte enters the detector cell and when signal is presented to the data processor. This response time has two principal contributions: the time for the detector cell to respond to the species, which may in turn be governed by a reaction rate in the cell (mass sensitive detectors), and the processing time which is determined by the amplifier time constant but involves the processing power of the electronic components used in assembly of the A/D converter. The first is a property of the detector and transient molecules and cannot be influenced except by changing detectors or sample. The time constant of the amplifier and choice of components is a matter of design. The detector amplifier is specified normally to filter or smooth high frequency background noise and for an FID electrometer it can range, depending on the age of the chromatograph, from about 500 ms (older) to about 50 ms (current). The requirement for fast chromatography is amplifier time constants in the region of 1 to 5 m and some manufacturers are beginning to address this requirement [14,15].

The properties of A/D converters surveyed by Kelly and Horlick were:

Aperture time	different sampling techniques, voltage/frequencies, dual slope A/D, require different times to measure the same data sample
Aperture jitter	random variations in the Aperture Time, these are small but technology must keep well ahead of narrowing peaks
Quantisation	minimum signal resolution and linear dynamic range
Sampling interval	the reciprocal of sampling frequency
Sampling period	the measured width of the peak

The advent of fast chromatography has brought a need to reappraise how the actual width of a narrow peak might impact on these quantities and affect the performance assumptions.

3.1. Aperture time

Converting an analog detector signal into a digital computer signal uses one of two principal A/D techniques, voltage/frequency conversion (VFC) or dual slope A/D conversion. Both involve charging a condenser, and dual slope A/D also involves the controlled and measured discharge of the same condenser. VFC is now considered to be slow and was more popular with earlier integrators than now. Later data processors use dual slope A/D.

The time interval for this signal conversion process is called the aperture time and its magnitude, a design feature, depends on the capacitance of the condenser. Ultimately the maximum sampling frequency of the A/D converter is set by the aperture time and by the RC time constant of this key condenser. It is a design assumption that aperture time, typically about 50 ms, is very much less than the peak width. If very narrow peaks are to be measured a smaller condenser must be used, and, as A/D converters are built onto a microchip, a new microchip is needed.

If aperture time is significantly large compared to peak width, the measurement of the peak must be synchronised with the operation of the aperture [16] or the measurement error varies. The random nature of retention times and regular sampling make such errors inevitable.

3.2. Aperture time jitter

Aperture time jitter is the variability of the A/D aperture. This is of concern when the sampling

interval is not negligible compared to the peak width. Normally peak widths are much broader than the fastest sampling frequencies of integrators, but as new and narrower peaks are measured with an older integrator perhaps, the problem can creep back.

3.3. Quantisation

The limit of quantisation is the smallest change in sample quantity to change the least significant bit in the A/D converter. This is not a high priority at the present time because micro-electronic technology is ahead of chromatographic requirements. Short columns and the limitations of injection technique mean that comparatively small quantities of sample are injected but these do not test the available limits of digital resolution. The use of 16 to 24 bit A/D conversion is more than adequate for the present. The linear dynamic range of these A/D chips is in the order of 10^4–10^5 and this is extended by auto-ranging.

Software writers of peak integration programs can effectively re-introduce quantisation error by truncating calculations too much or too soon in their programs.

3.4. Sampling interval

The key requirement of a data processor to measure very narrow GC peaks is fast enough data sampling; how fast is considered below. In other respects the measuring of narrow peaks is the same as the measuring of broad ones. There is a theoretical minimum number of samples for a measurement accuracy of say 0.1%, but data processors are able to produce a peak area from fewer samples than this minimum, because a measurement will take place provided the peak survives size and duration filters, and not through any consideration of peak base width and sampling frequency.

The sampling interval of a data processor is constant while a peak is being measured though it can be upgraded between peaks by ‘time to double’ commands. This constant sampling is convenient for a manufacturer but it imposes a mathematical limitation on the area measurement techniques which can be applied [17]. However, alternative sampling strategies such as random, variable or optimal sampling are scarcely worth implementing unless problems with experimental repeatability – i.e. predictable peak start and end, peak overlap and low signal/noise ratios are solved first; and when these problems are absent the standard integrator delivers adequate results.

Provision of fast data sampling by a manufacturer for the measurement of narrow capillary peaks will not necessarily require extra computer storage or higher processing speeds if analyses are short. But if these data processors are applied to normal chromatograms at the frequencies used for fast chromatography, they will quickly fill up their hard disks.

3.5. Sampling period and peak base width

Before integrators became memory based, i.e. with storage of data and hindsight calculations, the base width of a peak was implicitly defined by the data sampling. The beginning and end of peaks were the nearest edges of data samples and peak height/retention time were identified from the largest datum in between. Provided the data sampling interval is assumed to be very much narrower than the peak base this is not a source of great error. Peak boundaries were de facto synchronised with data sampling and the peak base width was an exact number of sample intervals but this was merely a consequence of no memory and not a physical reality, except by chance. The change from packed to capillary columns pressured the underlying assumption but it was a relatively easy thing to introduce integrators with higher sampling frequencies. When memory based integrators became available, they were able to store data until the end of the analysis and then perform calculations using data before and after events. Curve fitting and interpolation allowed peak boundaries and retentions to be located more accurately, between data points and not merely at the nearest datum edge. This effectively uncoupled the synchronisation of peak boundaries with data sampling, and in doing so changed the Newton Cotes formula for calculating measurement error from ‘closed’ to ‘open’, see below. The change to very narrow peaks as in FC and CEC again puts pressure on the assumption unless sampling frequency is again upgraded.

4. The mathematics of sampling

We can assess from standard numerical integration theory the number of samples to measure a peak at a specified accuracy.

4.1. Numerical integration of peaks

Chromatographic peaks are sampled and measured by Newton Cotes rules of which the trapezoidal rule and Simpson's rule are the most familiar. Newton Cotes rules involve the fitting of polynomial functions to regularly spaced data points. A cubic polynomial, the ''three eighths'' rule, is the highest in practice though there are no reports of its use to measure peaks, and the use of higher order polynomials does not necessarily lead to better accuracy [21]. The trapezoidal rule involves fitting a straight line between data points, Simpson's rule is more accurate because it fits a quadratic function to groups of three consecutive data points. Simpson's rule requires an odd number of data points per peak to provide an even number of data integrals for measurement, and ideally the peak maximum should be one of them. This is not too difficult to organise if sampling interval is small compared to peak width so that it can be assumed that peak start, maximum height and end synchronise with sampling. The greater accuracy of Simpson's rule is traded against fewer data points (for very narrow peaks), but this reduced sampling must not give rise to errors of under-sampling, loss of information or effective lack of synchronisation between sampling and peak boundary points.

4.2. Number of data samples/peak for accurate measurement

Kipiniak [18] considered the practical advantages of using Simpson's rule over the trapezoidal rule for measuring chromatographic peaks to be 'insignificant' but he was not measuring capillary peaks. The trapezoidal rule not only requires more data samples than Simpson's rule (well enough known) but those data processors which use it cannot sample the detector signal fast enough to provide accurate measures of the narrowest peaks, especially when the peaks are asymmetric.

The sampling error using Newton Cotes rules to measure peak area can be related to the number of data samples used in the measure as follows:

4.3. Trapezoidal rule

Suppose a continuous function, such as a peak shape, is sampled n times in equal time intervals and these samples are numerically integrated with the trapezoidal rule to measure the area. The difference between the true integral (area), and the measured integral is given by the composite error formula [19–21]:

$$\text{Error, } E = I_{\text{true}} - I_{\text{meas}} = (W_b^3/12n^2) \cdot |h''(t)| \qquad (6)$$

where W_b is the peak base width, i.e. the 'integration' limits; $|h''(t)|$ is the absolute value of the second derivative of the function with respect to time.

Note from Eq. (6) that the trapezoidal rule error is proportional to W_b^3; halving W_b decreases the error eight-fold. The cubic power of $(W_b)^3$ means that capillary peaks are predicted to have smaller measurement errors than 'packed column', i.e. broader, peaks of the same area provided that the peaks are sampled to the same extent. It also means that errors in measuring W_b are amplified by the same cubic power.

Measurement error is inversely proportional to the square of the number of samples/peak therefore the more measured samples/peak the smaller the error. Conversely, if the samples/peak become too few, if the peak is under-sampled, the error increases in a quadratic manner.

4.4. Simpson's rule

If these same n samples are numerically integrated using Simpson's rule to measure the peak, the difference between the true integral and the measured integral is given by a different composite error formula [19–21]:

$$\text{Error, } E = I_{\text{true}} - I_{\text{meas}} = (W_b^5/180n^4) \cdot |h''''(t)| \qquad (7)$$

where $|h''''(t)|$ is the absolute value of the fourth derivative of the function with respect to time.

This is the open error formula and assumes that the availability of data storage and curve fitting

techniques effectively uncouple peak boundaries and data sampling. It is not the closed formula of Kishimoto and Musha [22], where the denominator of 180 in Eq. (7) is halved. Before memory based integrators became available, peak boundaries were defined by data samples, there was no interpolation.

The dependence of Simpson's rule error in Eq. (7) on W_b^5 shows an even greater sensitivity to base width and the precise location of the peak boundaries, more so than the trapezoidal rule. The errors are inversely proportional to n^4 and fall quickly as the sampling frequency increases, but, conversely they expand just as quickly if peaks are under-sampled.

Eq. (6) is proportional to the second derivative and Eq. (7) to the fourth derivative of the peak shape function. All popular models of chromatographic peak shape involve exponential terms and asymmetry factors. The consequence of this is that the second and fourth derivatives, and hence the error, vary across the peak width and in consequence are sensitive to peak shape. Unless peak shape is constant and this is difficult to achieve, repeated measurements of the same peak at the same sampling frequency will provide similar measurements but with different sampling errors.

If the stable output of a peak simulator can be measured by a data processor at different, known, sampling frequencies, it is possible to test whether the data processor is sampling fast enough by doubling the sampling frequency and noting whether the improvement in accuracy, measured as a reduction in RSD, is negligible or not. It is also possible to plot sampling frequency versus measurement error to determine whether the data processor is using the trapezoidal rule or Simpson's rule, by determining whether the errors obey an inverse square or an inverse fourth law. Eqs. (6), (7) have different derivatives of E with respect to n. A good signal/noise ratio is required if these tests are made on real peaks.

4.5. Model for peak base width

In order to calculate the number of samples/peak required to produce a specified accuracy, base width is related to standard deviation, σ, and asymmetry. For the EMG peak model, asymmetry can be expressed in terms of τ, or τ/σ though σ is constant here:

Gaussian Peak $W_b = 6\sigma$ or 8σ (8)

EMG Peak [23,24]

$$W_b = 6\sqrt{(\sigma^2 + \tau^2)} \text{ or } 8\sqrt{(\sigma^2 + \tau^2)} \quad (9)$$

The theoretical boundaries of both peaks are $-\infty$ to $+\infty$, and are unmeasurable but selection of 6 times and 8 times the standard deviation of each peak sets a horizontal level above which the peak is measured and below which it is sacrificed as if lost in baseline noise. The area loss to a gaussian peak at $W_b = 6\sigma$ and 8σ is about 0.27% and 0.01% respectively – see Table 1. The area loss to an EMG peak at $W_b = 6\sqrt{(\sigma^2+\tau^2)}$ and $W_b = 8\sqrt{(\sigma^2+\tau^2)}$ is larger, it is about 0.6% and 0.08%. It would seem that the area of an EMG peak enclosed by 6 or 8 standard deviations is less than that enclosed by 6 or 8 standard deviations of a gaussian peak.

4.6. Evaluating the second and fourth derivatives

A gaussian peak shape is described by [24]:

$$h(t) := \frac{A}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{[-(t-t_r)^2]/2\sigma^2} \quad (10)$$

where $h(t)$ = the signal amplitude at a time t, A = peak area, t_r = peak retention time and σ is the

Table 1
Area recovery from integration of gaussian and EMG peaks [true peak area (all peaks), $I_{true} = 1.0000$]

Measured Peak Area:	τ/σ			
	0.0	1.0	2.0	3.0
Area at $W_b = 6$ SD	0.9973	0.9942	0.9942	0.9948
Area at $W_b = 8$ SD	0.9999	0.9993	0.9992	0.9992
Peak height	0.3989	0.3128	0.2376	0.1906

standard deviation. This function has second and fourth derivatives, $h''(t)$ and $h''''(t)$ given by:

$$h''(t) = H/\sigma^2 \cdot [(t - t_R)^2/\sigma^2 - 1] \cdot \exp[-(t - t_r)^2/2\sigma^2] \quad (11)$$

where H is maximum peak height. The value of the second derivative is not constant over the measured width of the function. It ranges from $0.446H/\sigma^2$ at $t = \pm\sqrt{3}\sigma$, to $-H/\sigma^2$ at $t = t_r$ and includes zero.

$$h''''(t) = H/\sigma^4 \cdot [3 - 6(t - t_R)^2/\sigma^2 + (t - t_r)^4/\sigma^4] \cdot \exp[-(t - t_r)^2/2\sigma^2] \quad (12)$$

The fourth derivative has five maxima and minima at $t = t_r$, and $(t - t_r)^2/\sigma^2 = (5 \pm \sqrt{10})$ and its values also includes zero. The largest absolute value of $h''''(t)$ is $3.815H/\sigma^4$. The largest errors, E_{max}, in Eqs. (6), (7) are derived from $|h''(t)|_{max}$ and $|h''''(t)|_{max}$.

Asymmetric peaks are not so easily described. The EMG model is used here, but the equation for this model has ''evolved''. Briefly, the original derivation of the EMG function was provided by Sternberg [2]. Changes to his equation were given by Foley and Dorsey [25] in 1984 and more corrections were made by Jeansonne and Foley [26] in 1991. Li [27] referred to ''slightly different forms'' of the equation in 1995 and submitted his version. It is Li's version, in error function format, which is used here to evaluate the second and fourth derivatives. The equation is:

$$h(t): = \left[\frac{A}{\sqrt{\pi}\cdot\tau}\cdot e^{[(1/2)\cdot(\sigma_g/\tau)^2 - [(t-t_g)]/\tau]}\right] \cdot \left[\frac{\sqrt{\pi}}{2}\left(\mathrm{erf}\left(\frac{t_g}{\sqrt{2}\cdot\sigma_g} + \frac{\sigma_g}{\sqrt{2}\cdot\tau}\right) + \mathrm{erf}\left(\frac{t_g}{\sqrt{2}\cdot\sigma_g} + \frac{\sigma_g}{\sqrt{2}\cdot\tau}\right)\right)\right] \quad (13)$$

In Eq. (13) previously used symbols have the same meaning, and: t_g = mean time of the gaussian component of the EMG function, σ_g = standard deviation of the gaussian component of the EMG function, τ = the EMG time constant, erf(..) = error function of (..).

Even this equation has its operational limits. For example, it does not converge on symmetry as $\tau/\sigma_g \rightarrow 0$. At $\tau/\sigma_g < 0.2$, the equation simply does not plot as a peak, while at $\tau/\sigma_g > 0.2$ it does. Similarly the function drops negatively at $t < -20$ if $t_g < 7$, though this is not important when applying it to chromatography. Eq. (13) is applicable in the range of τ/σ_g from 0.2 to 3.0. Asymmetries above this range are also described by Eq. (13) but the shapes are unacceptable and analysts would in practice seek to improve peaks with so much tailing.

Equations for the second and fourth derivatives cannot be derived as formulae from Eq. (13) but there are software packages such as Mathcad [28] which will evaluate $h''(t)_{max}$ and $h''''(t)_{max}$ and draw plots of the derivatives: an example is given in Appendix A. Like the derivatives of the gaussian function, the derivatives of the EMG have a range of values: positive, negative and zero.

Table 2 summarises the second and fourth derivatives used in the peak sampling calculations here.

This provides the information needed in Eqs. (6), (7) to evaluate n for the requirement that:

$$E_{max}/I_{true} \leq 1\% \text{ or } 0.1\% \quad (14)$$

Table 3 shows a summary of the minimum samples/peak to deliver these levels of accuracy, though it should always be remembered that they are based on E_{max}, the worst case. Figs. 2, 3 are graphical representations of Table 3; Fig. 2 shows the samples/peak required to deliver 1% accuracy, Fig. 3 shows the samples/peak for 0.1% accuracy. It is reassuring to note that although two completely different peaks' equations have been used, with different standard deviations and derivatives, the

Table 2
Largest numeric second and fourth derivatives of gaussian and EMG functions

τ/σ	0	0.25	0.50	1.00	1.50	2.00	2.50	3.00
$h''(t)$ max	−0.39894	−0.36805	−0.31219	−0.22143	−0.16541	−0.12993	−0.10606	−0.089128051
$h''''(t)$ max	+1.19683	+1.05235	+0.82761	+0.52868	+0.37417	+0.28572	+0.22974	+0.191530819

Table 3
Minimum number of data samples to measure gaussian and EMG peaks

Asymmetry	τ/σ							
	0	0.25	0.5	1.0	1.5	2.0	2.5	3.0
Minimum data samples for 1% error:								
Trapezoidal rule:								
Base width = 6×SD	26.8	26.9	28.0	33.6	41.8	51.1	61.0	71.2
Base width = 8×SD	41.3	41.5	43.1	51.7	64.3	78.7	94.0	109.7
Simpson's rule:								
Base width = 6×SD	8.5	8.5	8.9	10.7	13.2	16.2	19.4	22.6
Base width = 8×SD	12.1	12.2	12.7	15.3	19.0	23.2	27.7	32.4
Sampling ratio: Trapezoidal/Simpson's								
6 SD	3.16	3.16	3.15	3.15	3.15	3.16	3.15	3.15
8 SD	3.40	3.39	3.39	3.38	3.39	3.39	3.39	3.38
Minimum data samples for 0.1% error:								
Trapezoidal rule:								
Base width = 6×SD	84.7	85.2	88.6	106.2	132.1	161.7	193.0	225.2
Base width = 8×SD	130.5	131.1	136.4	163.5	203.4	249.0	297.2	346.8
Simpson's rule:								
Base width = 6×SD	15.1	15.2	15.8	19.0	23.6	28.8	34.4	40.2
Base width = 8×SD	21.6	21.7	22.7	27.2	33.7	41.3	49.3	57.6
Sampling ratio: Trapezoidal/Simpson's								
6 SD	5.62	5.62	5.61	5.60	5.61	5.61	5.61	5.60
8 SD	6.04	6.04	6.02	6.02	6.03	6.03	6.03	6.02

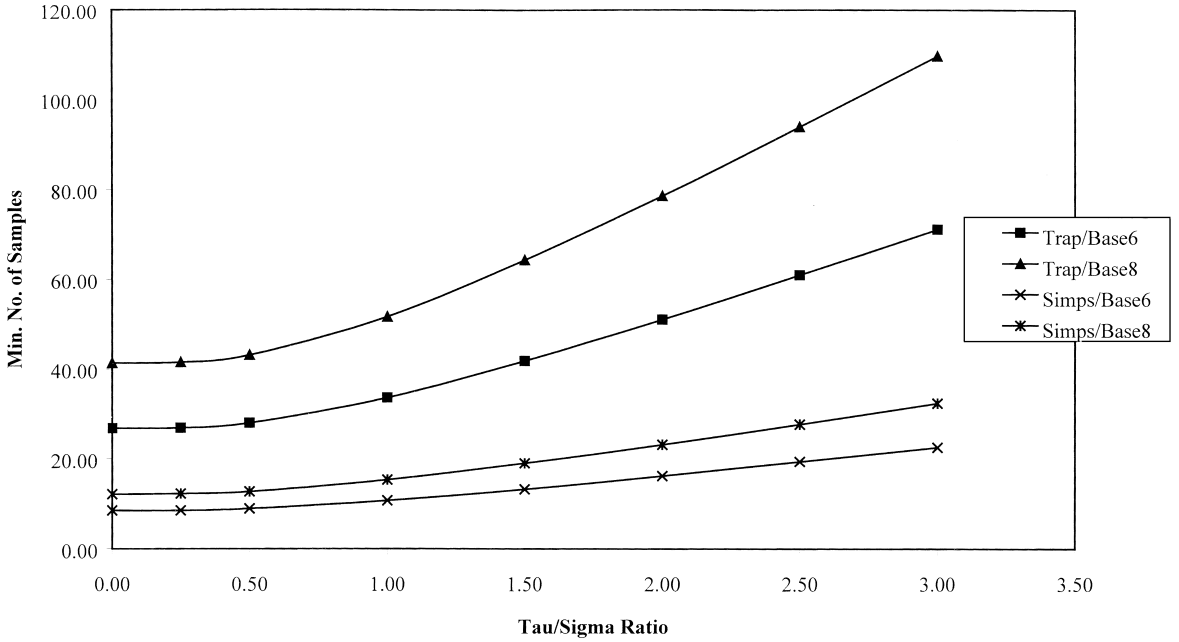


Fig. 2. Number of data samples for 1% error.

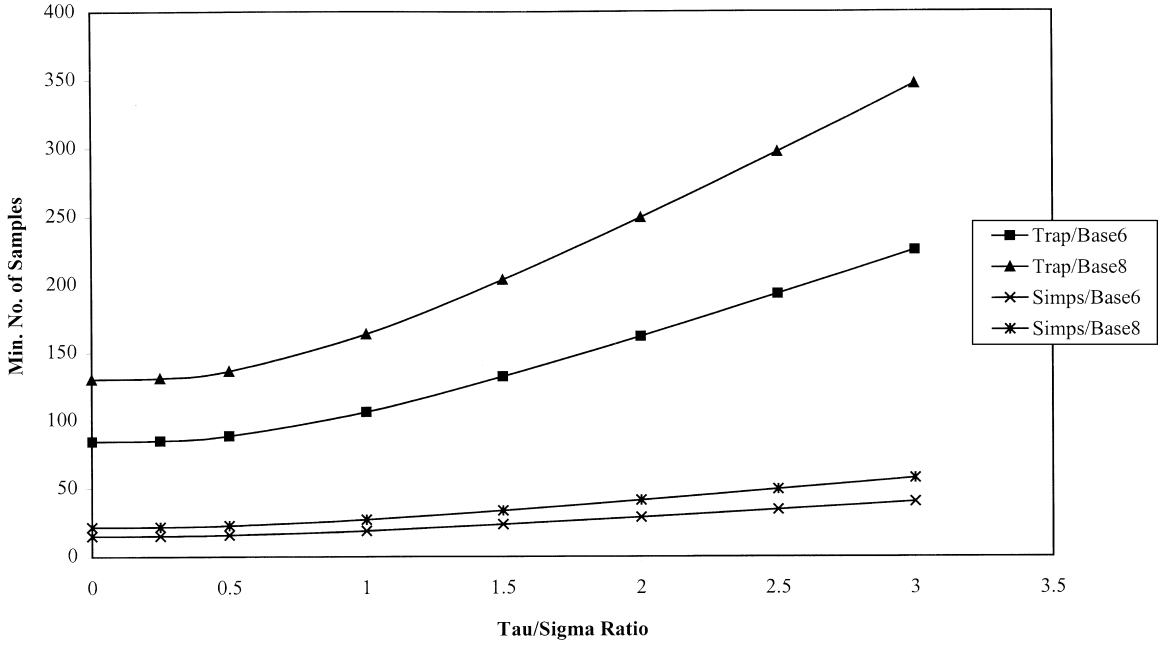


Fig. 3. Number of data samples for 0.1% error.

sampling figures for EMG peaks merge seamlessly on to the gaussian figures ($\tau/\sigma = 0$). The range of the figures however, quantify the penalty of asymmetry. The fact that fewer samples are required to measure peaks where base width is equal to $6\times$SD than where base width is $8\times$SD to provide the same measurement error does not mean that peaks on noisy baselines are easier to measure, merely that less of the total peak area is recovered. The sampling frequency is the same in both cases.

The assumption of ''about 20 to 30 samples/peak for accurate measurement'' can now be judged in context. The use of the trapezoidal rule to measure asymmetric peaks requires many samples/peak to achieve even 1% accuracy. If narrow capillary peaks, in this asymmetry range, of base width 0.01 s are to be measured by the trapezoidal rule at 0.1% accuracy, data sampling at 13 kHz to 35 kHz is required – this is into the radio frequency spectrum [29]. Simpson's rule is less demanding, 2 kHz to 6 kHz will provide the same accuracy. Peaks of 0.1 s base width (today's fast peaks) can be sampled at 200 Hz to 600 Hz depending on asymmetry.

Table 3 also compares the number of required data samples when measuring EMG peaks by each rule. The greater the accuracy required, the better Simpson's rule becomes. For 1% accuracy the trapezoidal rule needs about 3 times the number of samples that Simpson's rule needs, but for 0.1% accuracy the trapezoidal rule needs about 6 times as many. Simpson's rule is more efficient as well as more accurate and offers better protection against under-sampling than the trapezoidal rule. This is reflected in the spacing of the plots in Figs. 2, 3.

A survey of the those papers on peak measurement which include reports of the number of samples/peak used shows that about 20 to 30 samples/peaks are often used. The reasons behind these numbers are usually traced back to the papers of Kishimoto and Musha [22] and Chesler and Cram [30]. The conclusion to be drawn from Table 3 is that unless peaks are near symmetrical and measurement is known to have been made by Simpson's rule, 20 to 30 samples/peak may be gross under-sampling. Chromatographic data processors do not as a rule provide peak sampling information in their reports or indicate the possibility of under-sampling.

For a gaussian peak of base width 8σ, measured by Simpson's rule, Table 3 broadly agrees with Kishimoto and Musha although they measured height

rather than area, and measured error using the non-composite error formula [19–21].

From the theory, trapezoidal measurement of a symmetrical peak to 0.1% accuracy can require 130 or more data points. If the maximum sampling frequency of a data processor is 20 samples/s, the peak must have a base width in excess of 6.5 s. Asymmetric peaks require even more samples/peak than symmetrical ones; an asymmetric peak, $\tau/\sigma = 3$ requires nearly 350 data samples and its base width would have to be greater than 17 s. The same peak would only have to be 3 s wide to be measured to the same accuracy by Simpson's rule.

What is clear from Table 3 is that the measurement of very narrow capillary peaks requires Simpson's rule rather than the trapezoidal rule. The trapezoidal rule needs sampling frequencies that are not available in existing data processors. It is easier to provide better peak measurement by modifying the integration software than changing the sampling hardware, therefore the use of the trapezoidal rule to measure peaks is something that will die out if it has not done so already.

Assuming Simpson's rule is used to measure very narrow peaks, then from Table 3 it will require a sampling frequency of 5.76 kHz to measure an EMG peak of $\tau/\sigma = 3$ over a base width of $8(\sigma + \tau)$ with a sampling error of 0.1%. Chromatograms which are measured in minutes will be stored in files measured in megabytes; strategies for signal/noise enhancement will have to be reviewed. These data processors are not for routine chromatography. The trapezoidal rule would require a sampling frequency of 34.68 kHz!

5. Additional measurement diagnostics from the integrator

A chromatographic data processor should always provide as much information as possible about the measurements it makes so that the analyst knows what it has done. There are two important omissions in most commercial systems:

- There is no indication of what Newton Cotes summation rule is employed.
- There is no warning of under-sampling.

It is a weakness in data processors that a peak area is reported simply because it is big enough and wide enough to avoid being filtered on detection, and its measured size is larger than a pre-set minimum area or height and so it is not omitted from the report. This does not take into account whether the peak signal was sufficiently sampled which in turn involves the area summation rule.

It would be useful if data processors could provide a warning of under-sampling. This diagnostic would combine the actual number of (bunched) data samples used in a peak measurement, peak asymmetry, i.e. τ/σ or Asymmetry ratio [31], and the signal/noise ratio into a single conclusion.

This may be easier to request than organise, therefore a simpler diagnostic would be to report the actual number of data samples used in a peak measurement (not the same as the original sample frequency if ''time to double'' commands have been executed), the peak asymmetry (already available in most integrators) and the sampling frequency quoted in clear units of seconds or milliseconds. From this information the analyst can judge whether any measured peaks have been under-sampled. For the future it may transpire that a single dimensional measurement of peak asymmetry such as τ/σ or B/A ratio may not be sufficient to fully describe or monitor the shape of a three dimensional, dynamic zone which is the reality of an eluting peak.

Manufacturers do not publish information on their choice of Newton Cotes method for peak measurement; it should be included in the manual. It is clear from Table 3 that a measured peak might, or might not, have been under-sampled depending on how the area was calculated. Older integrators used the trapezoidal rule when most peaks were non-capillary, but methodology has changed; most peaks are now capillary and it is not clear in manuals whether existing data processors have switched to Simpson's rule (some may have always used it). Laboratories might be using a mixture of calculation methods in variously aged data processors, instruments with the same outward appearance and model number. Older integrators may still be in use measuring, and under-sampling, capillary peaks.

A Practical Note: under-sampling is a disqualifier of results and reports that indicate peak under-sampling will not be popular to users unless the sampling frequency of the integrator can then be increased to measure peaks without disqualification. The availa-

bility of this diagnostic might reveal the need to buy a new data processor with a faster sampling specification and thus threaten expenditure, but it would contribute to better chromatography.

6. Data processors and the errors of measurement algorithms

Data processors, characterised by their use of perpendiculars and tangents to separate overlapping peaks, are very good at their job provided that: signal/noise ratio of the detector signal is high, the peak zones are broad enough to measure, there is no significant overlap with other peaks, peaks are symmetrical or nearly so, and the baseline is flat.

If a laboratory integrator is provided with a stable signal from a peak generator, it will easily measure peak areas with a precision of 10^{-3}% [32].

When these conditions do not apply, measurement of real chromatographic peaks is vulnerable to errors of bias and imprecision, even when the data processor is programmed correctly. When the detector signal is not correctly aligned with the integrator's operating range, when this signal is sampled too slowly, when noise sensitivity is set too low (too insensitive), original data is lost from the detector signal and the problems multiply. Worst of all perhaps, is that under these adverse conditions, inaccurate area or height measurements are usually accompanied by accurate retention time measurements which can draw the uncritical into believing all of the results. The argument of whether area or height is the better measure rumbles on [33–35]. Area ought to be the better measure but there are many local instances where overlap, asymmetry and size make height a safer choice. Unfortunately those engaged in capillary zone electrophoresis may not have a choice because the effect of electromigration dispersion means there is 'often no proportionality between peak height and injected amount' [36].

Further limitations of integration algorithms are exposed when confronted by the other major problems of chromatography which are the following.

6.1. Baseline noise

Problems with baseline noise arise when the signal/noise ratio is low, as in trace analyses, because peak boundaries are obscured, introducing uncertainty into baseline allocation and subsequent peak measurements. Attempts to suppress the noise risk distorting the required peak unless there is some unique feature of the noise to target [37]. Noise at peak maxima can trigger peak splitting. Signal/noise ratio determines the minimum detectable quantity – usually set at twice the background noise level if you know where on the baseline to look, or three times the background if you do not.

There are three basic types of noise encountered in chromatography: (1) Random, high frequency noise and spikes – much narrower than peaks; (2) Noise that looks like small unexpected peaks – and may well be; (3) Wandering baseline or drift – too slow and poorly shaped to be confused with peaks.

Some of these types have characteristics that allow them to be successfully suppressed, others have not.

Random or white noise is removed by the time averaging of the signal in three ways: by RC time constants in the detector and electronics, by sampling the detector signal which time-averages the signal over the sampling interval and by using moving window smoothing techniques during the subsequent processing. This is an area where the analyst has little or no control over the applied processing as the algorithms are provided by the manufacturer. At best there is some algorithm selection such as setting the peak width or slope sensitivity parameter.

Most chromatographers are surprised when they first view the detector signal of a peak on an oscilloscope before it passes into the detector electronics, and compare it to the same peak on a chromatogram. The quantity of background noise and narrower peak width displayed by the oscilloscope is significantly different because the time constant of the oscilloscope is much less than the combined time constant of detector electronics and data processor, and the smoothing algorithms of the data processor have been by-passed.

Baseline "rubble" or small unexpected peaks may simply be a real part of a complex sample matrix. As unwanted peaks they are noise, and if in their abundance they obscure the baseline position, they hinder the accurate measurement of wanted peaks. Electronic or mathematical filtering of these unwanted matrix peaks would also interfere with the

peaks to be measured and is therefore inappropriate. Such background is best removed by sample and column conditioning or by selective detection. The data processor can differentiate and remove the smallest peaks by applying a minimum size criterion.

An unfortunate consequence of time averaging high frequency noise is that much of it is not entirely random and smoothing makes it look like small peaks which are then measured. Their appearance on the baseline may be interpreted as a need for column re-conditioning.

Drift or wandering baseline has no usable feature for characterisation; it indicates an unstable experiment. Drift is supposed to be removed during experimental setup. If it appears during an experiment it is a warning of some emerging problem. In consequence when drift appears data processors have no tools to deal with it other than post run recovery of information. The analyses immediately before must be inspected for errors and their results have lower confidence limits until the cause of the drift problem has been discovered and removed.

6.2. Separating overlapping peaks [38]

The use of perpendiculars and tangents to separate overlapping peaks was shown to be inaccurate by Westerberg [39] in 1969, and it was anecdotal before that. Overlapping peaks can be separated by a perpendicular dropped from the valley point between them only if the peaks are symmetrical, if they have the same shape and size and they have the same response factors within that experiment. Measurement errors enter in all other conditions and the valley bottom is not then the correct place to drop a perpendicular.

Peak overlap does not reduce the signal, it masks it. Overlapping peaks add linearly and if the sampling frequency is sufficient to measure them individually it is sufficient to measure them collectively. The accuracy of measurement of overlapping peaks is determined by the method of separating the peaks and not by the sampling frequency provided the sampling frequency is fast enough.

Attempts to correct perpendicular errors with correction coefficients have generally failed because coefficients imply a peak shape, are geometric in character and cannot account for different response factors. Non-repeatability of chromatograms and variability of valley height and valley position hinder the application of correction coefficients.

If a valley is not the correct place to cleave two overlapping peaks to achieve accurate and separate areas, but noting that there is a correct though undefinable cleavage point somewhere between the two peaks, the consequence of separating peaks at the valley is to transfer the area enclosed between the valley and the correct cleavage point from one peak to the other. The area of one peak is increased by $+\delta A$, the area of the other is decreased by $-\delta A$. The sum of the two measurement errors, $[+\delta A+(-\delta A)]$, is zero. This idea can be extended to groups of peaks: provided that the baseline is drawn correctly beneath a group, the sum of all the area transfers within a group of n peaks, caused by perpendicular separation at the valley point, is zero. The measured total area is accurate; only the individual areas are inaccurate:

$$\Sigma_{n}(\delta A_{i}) = 0 \tag{15}$$

from which we can extract the measurement error of any one peak, x, within the group as:

$$\delta A_{x} = -\Sigma_{n-1}(\delta A_{i})\ (i \neq x) \tag{16}$$

which means that the measurement error in any peak in a group is equal and opposite in sign to the sum of all the measurement errors of the other peaks in the group. Depending on the population of the group, the degree of overlap, and the peak sizes and asymmetries, this can amount to a very large error and make the measurement of peaks within a group effectively worthless. The object of method development is therefore to make overlapping groups contain as few member peaks as possible.

6.3. Tangent skimming

Measurement of small and narrow peaks on the tails of larger peaks is made by skimming a tangent under the smaller peak and measuring its area above this tangent. The height ratio to trigger this type of measurement is usually about 10:1 but it is an arbitrary ratio, some manufacturers include a width ratio criterion to ensure only narrower peaks are skimmed. Such measure is accurate if the smaller

peak is very much smaller and narrower than the host, but accuracy also depends on the position of the smaller peak on the tail [24], accuracy improves as the smaller peak moves down the tail, as resolution increases.

The transition from perpendicular separation to tangent skimming creates a discontinuity in the measurements of marginally sized peaks, but data processors offer the analyst the means to consistently enforce one measure or the other.

6.4. Asymmetry

Asymmetry is not a direct problem to a data processor because measurement of peaks by signal sampling and summation makes no assumptions of shape, though according to Table 3, more samples/asymmetric peaks are required to deliver the same level of accuracy. An asymmetric peak can, in principle, be measured as well as a symmetrical one. Problems of asymmetry are caused when the spreading peak base reaches other peaks and introduces or extends the problems of overlap. Noise on a shallow gradient peak tail can fool an integrator into ending peak measurement too early.

Variable asymmetry within an overlapping group frustrates simple modelling [40]. If asymmetry was constant within a chromatogram, shape factors could be applied to the peaks to correct for the overlap errors, but asymmetry is rarely constant and such a simplification becomes impossible. Unlike area correction coefficients, shape factors are independent of response factors.

Peak tailing creates a systematic error: when overlapping asymmetric peaks are separated by perpendiculars, the tail area of the earlier peak is added to the next peak and exaggerates its area. If several asymmetric peaks overlap, each peak, except the first, will receive the tail of its predecessor and have its area enlarged. This will be offset, to an extent depending on peak size, by the transfer of its own tail to its follower – except for the last peak which has no follower to donate to, and this can create an under-estimate of the first peak area in the group and an over-estimate of the last. The measurement error is systematically transferred along the group. If the tail of a peak extends past the maximum of its next neighbour, it lifts the neighbour and makes height measurement inaccurate.

6.5. Baseline allocation

Data processors draw linear baselines beneath peaks to complete the boundary for measurement. This is acceptable when the baseline is flat, but gradient chromatography often creates non-linear baselines and in such experiments it may be perfectly obvious that the baseline, i.e. the signal created by the analysis without sample introduction, should be curved. Baseline events such as valve switching spikes, negative detector signals and stray peaks from older analyses add to the problems of baseline allocation.

Tangent peaks are sometimes skimmed from host peaks by curve fitting an exponential baseline. This is acceptable in principle but, because of a lack of detail in manuals, the analyst is never quite sure how the exponential baseline copes with variations in peak shape in repeat experiments.

7. Calibration

Inaccurate measurements are avoided by calibration and the use of response factors, but these response factors incorporate a degree of empirical compensation; they are not independent of the particular chromatograph or analysis. In other words the response factors contain an embedded instrument factor and are analysis specific. As long as this embedded instrument factor is constant, which means that provided the chromatogram has the same profile, same peak sizes and positions, same degree of overlap and same baseline shape, and provided that calibration has spanned the whole measuring range, the results are credible. Unfortunately for the analyst, the only means of knowing that this instrument factor has drifted is when the results have become noticeably wrong and this means re-calibration of the system and checking some old results. In addition this empirical accuracy does depend on the integrator measuring peaks accurately which is not the case if they have been under-sampled or the integrator has not been programmed correctly in general.

There are no international standard chromatograms against which to calibrate integrators. Standard solutions can be analysed of course, but data processor

and chromatograph form a closed loop which conceals systematic errors and the precision of measurement of such chromatograms is only about 0.1% on a good day. While this precision may be sufficient for a particular assay it is not really impressive compared against best scientific measurements. If international standards of mass or voltage were only measured to 0.1%, science would have achieved much less than it has.

Electronic signal generators deliver a more precise signal (RSD about 10^{-3}%) [24] and can be used to calibrate the data processor independently of the chromatograph, but they, in turn, are not independently calibrated against mass or coulombs, and they do not incorporate response factors though these can be added as empirical inputs.

Nevertheless, and this is a recurring theme in the practical measurement of chromatograms, proper method development, knowledgeable use of the data processor, careful calibration and vigilant conduct of the experiment will compensate for the deficiencies of the chromatograph and data processor, and produce accurate assays.

8. Improving data processors

There is a large body of theory and practice on signal enhancement and resolution concerning the extraction of information from data (for a start see Ref. [37]). Techniques involving peak modelling and signal conditioning are widely practised in other branches of science, but they have made only modest impact on chromatography. Compared to the enhancement of satellite photographs from space or the resolution of radar images, the measurement of overlapping peaks is unsophisticated. Chromatographic data processors continue to use perpendiculars and tangents because there is still nothing better, even signal/noise enhancement of asymmetric peaks is problematic. Integrators are one of the few measuring devices that do not include measurement accuracy in their specification.

There is no lack of ideas, application or expertise for the development of more sophisticated data processors. Statistical mathematics, physics, electronics, spectroscopy, have all provided a number of techniques which have been tried in chromatography laboratories with some success – and that is the problem. All of the techniques are successful in some circumstances but not others. Analysts cannot depend on a data processor that measures some peaks but not others, or subtly worse, that measures all peaks with varying and unknown accuracy. Tangents, perpendiculars and linear baselines may be crude but they are simple to understand and predictable within their limits. Calibration defines those limits.

More sophisticated techniques of signal measurement fail in chromatography for a number of theoretical and practical reasons.

8.1. The dynamic nature of chromatography

Consecutive analyses of the same sample are unlikely to generate the same data set of measurements. The dynamic nature of chromatography, ageing of columns, samples with undetected or immobile components conspire to create an experiment where the rules of peak shape formation change from one experiment to the next. Spectroscopy is static by comparison and methods of peak modelling and deconvolution work better there.

8.2. Not enough information in FID and UV detection signals

FID in GC and UV in HPLC are the most widely used detections systems but they are essentially ''bulk'' detectors and their signals do not carry enough specific information to allow peaks to be deconvolved mathematically. If deconvolution is to succeed it will do so with detectors that carry more information, for example diode array or GC–MS, or in general, detectors that allow both selective and non-selective detection.

If an FID system (say) carried sufficient information for deconvolution then a single peak could, in principle, be separated into its components by mathematics alone and there would be no need for a column. The deconvolution algorithms could be tested for functionality by making them resolve a single species into its isomers; the correct isomer count and abundance figures would be a proof. These algorithms would not confuse two peaks with the same retention, they would spot the difference (i.e. identify the analytes) and provide different isomer counts for each. Clearly this is impossible, the FID

and UV detection do not have the means to identify molecules. Unless information rich detector signals are used, all except the simplest attempts at mathematical deconvolution will fail, and they have.

8.3. Knowing the number of components in a sample

It is very difficult to know the number of components in a sample. In some instances,‘‘ number’’ is a point of semantics – are isomers different components? In one analysis yes, in another no, it depends on the purpose of the analysis. Peaks are distribution/dispersion phenomena and while it might be possible to deconvolve them by fourier transformation into a set of components, each peak or group will deconvolve into many equally valid sets. The analyst must know how to select the ‘‘correct one’’ and then the results are not necessarily accurate.

Fellinger has recently shown [41] that it is impossible to conserve an asymmetric peak shape during fourier transform deconvolution. The variable asymmetry which is common in overlapping groups makes it impossible to select a single transfer function and so one of the most important techniques of peak/peak deconvolution is effectively hobbled. Economou et al. [42] tried fourier transform on peak height rather than area and found that, ‘‘under suitable circumstances’’, they too could deconvolve symmetrical peaks but not asymmetric ones.

For the task of signal/noise enhancement, a separate issue to peak/peak resolution, the general exponential character of chromatographic peaks means that they include a wide range of fourier components. Some of these fourier peak frequencies match those of the background noise. Schemes to remove noise inevitably remove matching peak frequencies and so distort the original peaks; the removal of these frequencies may increase or decrease the reconstructed peak area depending on whether the filtered frequencies would have recombined destructively or constructively. Smeller et al. [43] warn of other effects of noise on deconvolving, and over-deconvolving, symmetrical peaks. The noise can be made worse and side lobes appear on the peak which might be mistaken for and measured as other peaks.

The prime target of peak/peak resolution in FID and UV detection signals has given way to the easier task of signal/noise enhancement. Provided they are big enough, peaks are easier to distinguish from baseline than from each other.

While it is unlikely that all components (however defined) in research or environmental samples are known, some production and quality control chemists may legitimately argue that they do know what is in their samples and their experiments have good signal/noise ratios. In such cases a deconvolution based data processor might be possible: deconvolution being guided by the additional information, including peak shape that the analyst can provide. But this data processor will be laboratory specific and therefore unattractive for any manufacturer to provide. It has to be a very special analysis and a significant improvement in accuracy to persuade a laboratory to switch from the conventional approach of developing methods tolerant of commercial data processors. If the need is great enough, analysts must be prepared to act alone and develop their own data processor.

8.4. Suitable peak models

Chromatographic peak shape is determined before analyte reaches the detector. It is determined by interaction between analyte and column, by dispersion during elution and by extra-column broadening effects. The sensitivity of the detector determines peak size but not shape as long as the detector time constant is small and the reaction rates in mass selective detectors are much faster than response times. Perfectly pure species detected in isolation will produce a range of peak shapes on a range of columns. It is not surprising that one peak model (gaussian, EMG, etc.) will not fit all peaks. What is surprising, is that a few models can be tweaked to fit so many peaks. Data processors which rely on deconvolution methods will need to draw from a range of peak shapes and have a method of selecting and matching them to each peak. Even then, it will always be possible to select the wrong model or over-deconvolve and produce false positive results. Dimandja et al. [44] used the EMG function and maximum entropy methods to assay the Venus

atmosphere from data gathered by a Pioneer spaceflight and were able to produce spurious peaks which had ''no physical meaning''. They concluded that the EMG function was not a suitable model for N_2, but they only found out by spotting a false positive they could easily identify. This was in a known four component mixture: N_2/Ar/CO/Freon, their Fig. 5 [44]. In a more complex chromatogram, the false result might not have been so easily seen.

8.5. Lack of a traceable standard chromatogram

Chromatography is comparative science rather than absolute. If two samples produce the same chromatogram under identical conditions they are judged to be the same sample. To become an absolute science, and that is the direction in which regulation and legislation seems to be driving it, chromatography requires a standard chromatogram traceable back to international standards of mass, not voltage. This chromatogram will be used to calibrate integrators and data processors. There will be difficulty in agreeing upon a standard chromatogram because real chromatograms are so diverse. None except the simple would attract wide support but this will introduce the problem of how to relate a simple standard to laboratory analyses.

Peak generators have stable output signals and could be developed into a standard device, but they need to be made traceable to mass. They can then calibrate integrator counts to mass, though problems with overlap of peaks and response factors would remain.

8.6. Cost and complexity of alternative data processing techniques

The various theories and methods of deconvolution are difficult to understand in detail and unless data processors are to become black boxes with password protection controlled by the laboratory manager, mis-application of methods is inevitable. Training analysts and technicians to use complex data processors will pose a bigger problem than training them to use current models which employ perpendiculars and tangents.

Even if accurate deconvolution techniques were to be developed they would not necessarily be adopted. Standard experiments would have to be redeveloped with new (information rich) detectors [24], re-calibrated and re-validated because empirical response factor values depend on the method of separating overlapping peaks as well as on detector sensitivity. The cost of implementing new integration techniques will extend far beyond the purchase cost of the integrator. Some laboratories would simply opt not to use the new data processors on the grounds that it is too expensive to introduce them and the old methods are good enough.

Such considerations must cause any manufacturer to pause before developing this new breed. Will there be enough of a market to provide a return on investment? Will technical support teams be overwhelmed? For the present we are stuck with perpendiculars and tangents.

Acknowledgements

I would like to extend my thanks and appreciation to all those who have supplied me with information and details for the preparation of this article, including in alphabetical order: Gunther Braun of Shimadzu GmbH, Germany; Ian Dale of Gynkotek UK; Professor Walter Jennings of J & W Scientific; Bob Johnson of Varian UK; Brian Keenan of Bio-Rad UK; Martin Perry of Aston Scientific UK; Ken Riley of Hewlett-Packard UK; Gerard Sharp of SGE International, Australia; Keith Wright of SGE UK.

Appendix A. Derivatives of EMG based on error function

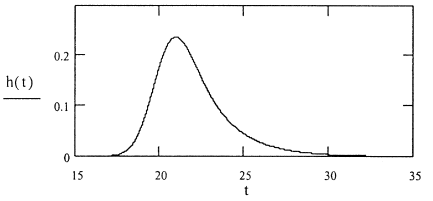
Parameters $A:=1$, $\tau=2$, $\sigma g:=1$, Gaussian retention, $tg:=20$. Range of time variables, $t:=15$, 15.011 . . 35. Define

$$I(t):=\frac{\sqrt{\pi}}{2}\left(\operatorname{erf}\left(\frac{tg}{\sqrt{2}\cdot\sigma g}+\frac{\gamma g}{\sqrt{2}\cdot\tau}\right)+\operatorname{erf}\left(\frac{t-tg}{\sqrt{2}\cdot\sigma g}-\frac{\sigma g}{\sqrt{2}\cdot\tau}\right)\right)$$

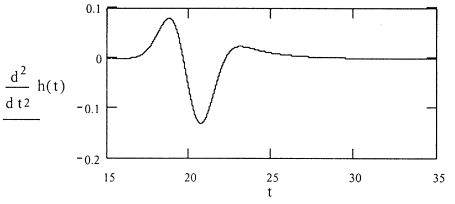
EMG function:

$$h(t) := \left[\frac{A}{\sqrt{\pi} \cdot \tau} \cdot e^{(1/2)\cdot(\sigma g/\tau)^2-(t-tg)/\tau]} \right] \cdot I(t)$$

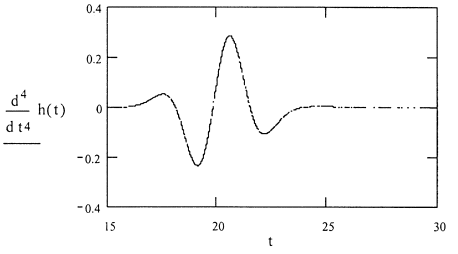
Standard deviation: $\sigma = \sqrt{(\sigma g^2 + \tau^2)} =$ 2.2360679775. Peak height = 0.2376365698.



Second derivative: Upper limit = 0.0794898423. Lower limit = −0.1299267757.



Fourth derivative: Upper limit = 0.2857162498. Lower limit = −0.2328309948.



References

[1] W.B. Furman, T.P. Layloff, R.F. Tetzlaff, J. AOAC Int. 77 (1994) 1314.
[2] J.C. Sternberg, Adv Chromatogr. 2 (1966), Ch. 6.
[3] M.H.J. van Rijswick, Philips Research Reports Supplements, No. 7, Centrex Publishing Co., Eindhoven, 1974.
[4] M.J.E. Golay, in: D.H. Destry (Ed.), Gas Chromatography, Butterworths, London, 1958, p. 36.
[5] G. Gaspar, R. Annino, C. Vidal-Madjar, G. Guiochon, Anal. Chem. 50 (1978) 1512.
[6] J.R. Torres-Lapasio, J.J. Baeza-Baeza, M.C. Garcia-Alvares-Coque, Anal. Chem. 69 (1997) 3822.
[7] R. Annino, J. High Resolut. Chromatogr. 19 (1996) 285.
[8] M.A. Klemp, M.L. Akard, R.D. Sacks, Anal. Chem. 65 (1993) 2516.
[9] R. Annino, J. Chromatogr. A 678 (1994) 279.
[10] J. Beens, H. Boelens, R. Tijssen, J. Blomberg, J. High Resolut. Chrom. 21 (1998) 47.
[11] J.P. Kutter, S.C. Jacobson, N. Matsubara, J.M. Ramsey, Anal. Chem. 70 (1998) 3291.
[12] R. Dadoo, R.N. Zare, C. Yan, D.S. Anex, Anal. Chem. 70 (1998) 4782.
[13] P.C. Kelly, G. Horlick, Anal. Chem. 45 (1973) 518.
[14] D. Coe, C. Hodges, M. Klemp, H. Kern, presented at the 16th Poster Publication, International Symposium On Capillary Chromatography, Riva del Garda, 27–30 September, 1994.
[15] M Series Micro GC Sales Literature, Hewlett-Packard.
[16] D.T. Rossi, J. Chromatogr. Sci. 26 (1988) 101.
[17] G. Fox, http://www.npac.syr.edu/users/gcf/CPS615NI95/node4.html, Numerical Integration Module. Also, internet Keyword: Newton Cotes (exact phrase).
[18] W. Kipiniak, J. Chromatogr. Sci. 19 (1981) 332.
[19] R.L. Burden, J.D. Faires, Numerical Analysis, 6th ed, Brooks/Cole, Pacific Grove, CA, 1997.
[20] E. Kreyszig, Advanced Engineering Mathematics, 7th ed, Wiley, 1993.
[21] C. Woodford, C. Phillips, Numerical Methods with Worked Examples, Chapman & Hall, London, 1997.
[22] K. Kishimoto, S. Musha, J. Chromatogr. Sci. 9 (1971) 608.
[23] J.P. Foley, J.G. Dorsey, Anal. Chem. 55 (1983) 730.
[24] N. Dyson, Chromatographic Integration Methods, 2nd ed, Royal Soc. of Chemistry Monographs, Cambridge, 1998.
[25] J.P. Foley, J.G. Dorsey, J. Chromatogr. Sci. 22 (1984) 40.
[26] M.S. Jeansonne, J.P. Foley, J. Chromatogr. Sci. 29 (1991) 258.
[27] J. Li, J. Chromatogr. Sci. 29 (1995) 568.
[28] Mathcad8 Professional Software, Mathsoft Inc., Cambridge, MA.
[29] Chambers Science and Technology Dictionary, Revised Edition, W&R Chambers Ltd, Edinburgh, 1974.
[30] S.N. Chesler, S.P. Cram, Anal. Chem. 43 (1971) 1922.
[31] J.J. Kirkland, W.W. Yau, H.J. Stoklosa, C.H. Dilks Jr., J. Chromatogr. Sci. 15 (1977) 1735.
[32] N. Dyson, Int. Lab. 22, No. 6 (1992) 38.
[33] V.R. Meyer, LC·GC Int. 7 (1994) 94.
[34] L.R. Snyder, J. Chromatogr. Sci. 10 (1972) 200.
[35] R.A. Keller, J. Chromatogr. Sci. 11 (1973) 223.
[36] H. Poppe, Adv. Chromatogr. 38 (1998), Ch. 6.
[37] E.R. Davies, Electronics, Noise and Signal Recovery, Academic Press, London, 1993.
[38] V.R. Meyer, Adv. Chromatogr. 35 (1998), Ch. 8.
[39] A.W. Westerberg, Anal. Chem. 41 (1969) 1770.
[40] D.W. Kirmse, A.W. Westerberg, Anal. Chem. 43 (1971) 1035.
[41] A. Fellinger, Anal. Chem. 66 (1994) 3066.
[42] A. Economou, P.R. Fielden, A.J. Packham, Analyst 121 (1996) 97.
[43] L. Smeller, K. Goossens, K. Heremans, Appl. Spectros. 49 (1995) 1538.
[44] J.M.D. Dimandja, M. Kaljurand, J.B. Phillips, J. Valentin, Anal. Chim. Acta 371 (1998) 1.