Journal of Chromatography, 550 (1991) 51–69
Elsevier Science Publishers B.V., Amsterdam

CHROMSYMP. 2256

Resolution enhancement of chromatographic data

Considerations in achieving super-resolution with the constrained iterative relaxation method

MARK R. SCHURE
Computer Applications Research, Rohm and Haas Company, 727 Norristown Road, Spring House, PA 19477 (USA)

ABSTRACT

The consequences of numerical resolution enhancement in column chromatography are examined with respect to the increase in the number of peaks that will become recognizable for analysis as deconvolution removes peak fusion and artificially extends the peak capacity of the chromatographic system. Using computer-generated chromatograms and numerically deconvolving with the constrained iterative relaxation method (CIRM), the results suggest that under ideal conditions deconvolution lowers the average resolution necessary for singlet peak discrimination to *ca.* 0.352. This decreased resolution limit allows the number of recognizable peaks, for fairly saturated synthetic chromatograms, to increase by *ca.* 50% when the ratio of the true number of components to the peak capacity equals 1.5.

INTRODUCTION

One of the most important advancements in separation science over the last two decades has been the development and application of high-resolution chromatographic methods aimed at resolving the components of complex mixtures of biological and environmental origin. Both capillary gas chromatography (GC) and high-performance liquid chromatography (HPLC) now have the performance to achieve high-resolution separations of complex samples that a short time ago would have seemed beyond the reaches of these techniques. The requirements for stringent quality control in biotechnology, controlled substance testing, forensic applications and environmental monitoring have placed even more demands on these techniques and consequently the use of very high-resolution columns, which are capable of generating over 10^5 theoretical plates, is expanding rapidly.

Although column technology will no doubt evolve in all application areas, there are a few applications where column technology has yet to make baseline separations a commonplace occurrence. In these applications, the components of the mixture often elute at such close retention times that individual separation and determination are difficult; however, in some instances the analyst can trap the peak and use another column of different selectivity. Alternatively, the operating conditions can be changed

0021-9673/91/$03.50

to run the experiment under the highest possible resolution (by changing the column temperature and flow velocity) or to use a capillary column instead of a packed column. When available, a multi-channel detector such as a mass spectrometer (GC) or UV array detector (LC) can be used to aid in the analysis of a mixture when peaks are not well resolved. This scenario, of course, requires that the analyst has recognized that peak fusion is present and that the proper instrumentation is available for this type of analysis. It would be most convenient under these situations if the analyst could somehow increase the resolution of the column to look for other closely eluting peaks. This desire for increased resolution will probably continue until column efficiencies are high enough to resolve every component or until resolution enhancement through multi-channel detectors and computer methods is routinely available.

A number of recent driving forces have created a situation where certain aspects of numerical resolution enhancement need to be re-examined, especially with regard to the single-channel detector most commonly used in chromatographic systems. These driving forces include (1) the realization through chromatographic theory [1–13] that for samples of complex origin the resolving power needed for baseline separations has been previously underestimated, (2) the need for higher resolving power in applications where components of minor concentration may play a major role in determining some property of the sample, (3) the recent production of numerical algorithms which are capable of high-resolution deconvolution and are now, owing to the increased computational performance of laboratory computers, feasible for use in laboratory applications and (4) the renewed interest in the flame ionization detector for HPLC, an inherently single-channel detector. Not only does high resolution help the analyst in the proper assessment of peak purity, important when fractions are collected for chemical analysis or when a "hyphenated" method of analysis is used; maximum resolution is of the utmost importance when automated method development is used.

This paper addresses the question of the extent to which in theory numerical deconvolution can offer quantitative resolution enhancement for complex chromatograms. Towards this goal, one of the best numerical methods of deconvolution, namely the constrained iterative relaxation method (CIRM), is used here to examine the limit of resolution enhancement. In addition, some aspects of deconvolution implementation are discussed.

THEORY

Potential benefits of deconvolution

The broadening of chromatographic peaks, as is well known and as demonstrated in Fig. 1, causes the recognizable loss of individual components through the fusing together of closely spaced peaks. In Fig. 1, the retention time and amplitude sequence is constant between synthetic chromatograms. Only σ, the standard deviation of the Gaussian peak shapes, is varied between the chromatograms. Unless shoulders or inflection points are found in the peak shape, it is often impossible for the analyst with a single-channel detector to determine if the peak is pure or composed of two or more components. The mathematical description of this loss of identity of pure peaks was recently described [1–13] for complex chromatograms possessing random retention times (additional assumptions of this approach are contained in the references). In this

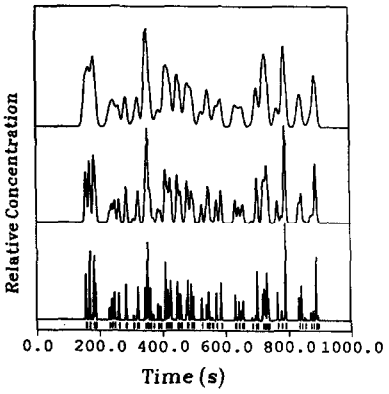

Fig. 1. Synthetic chromatograms of 80 Gaussian peaks generated with identical Poisson-distributed retention times, uniformly-distributed random peak heights (between 0.05 and 1), constant σ and no noise. Bottom, $\sigma = 1$ s; middle, $\sigma = 3$ s; top, $\sigma = 6$ s. Retention times are indicated by the bars below the bottom chromatogram.

theory the number of recognizable peaks, p, is related to the number of true components, m, via

$$p = m\exp\left(-\frac{m}{n_c}\right) \tag{1}$$

where n_c is the peak capacity [14], which is defined as the maximum number of component peaks which can be uniformly packed into a chromatographic elution profile at a stated resolution. The peak capacity, n_c, for situations such as gradient elution chromatography and programmed-temperature gas chromatography, where the peak variance, σ^2, is considered to be constant throughout the chromatographic elution profile, is given by [2,14]

$$n_c = \frac{t_{max} - t_0}{4\sigma R_s^{\ddagger}} \tag{2}$$

where t_{max} and t_0 denote the time range of the elution profile and $R_s^{\ddagger}$ is the stated critical resolution, which has been estimated for this statistical model to be 0.5 [3,11]. Other studies have suggested alternative values for the critical resolution, *e.g.*, 0.8 [9] and 0.71 [10]. For isocratic LC or constant-temperature GC, where it is assumed that the peak variance is not constant but rather increases with time [14],

$$n_c = \frac{\sqrt{N_{av}}}{4R_s^{\ddagger}} \ln\left(\frac{t_{max}}{t_0}\right) \tag{3}$$

where N_{av} is the average number of theoretical plates.

Eqns. 1–3 allow the extent of deconvolution to be related to the number of observable peaks by the realization that deconvolution increases the peak capacity by synthetically lowering $R_s^{\ddagger}$. An expression is now obtained whereby the number of

peaks visible after deconvolution of the chromatogram, p_d, is ratioed to the number of peaks visible in the native chromatogram, p. From eqn. 1:

$$\frac{p_d}{p} = \frac{m \exp\left(-\frac{mR_s^d}{n_c'}\right)}{m \exp\left(-\frac{mR_s^\ddagger}{n_c'}\right)} \tag{4}$$

where R_s^d is the critical resolution of the deconvolution technique and n_c' is the peak capacity with unit resolution, *i.e.*, $n_c' = n_c R_s$. Our use of the term resolution here is in keeping with the standard definition of resolution, $R_s = \Delta t / 4\sigma$, where Δt is the difference in retention time, t_r, between two peaks whose Gaussian standard deviations are both equal to σ. Rearranging eqn. 4 yields

$$\frac{p_d}{p} = \exp\left[\frac{m}{n_c} \cdot \left(1 - \frac{R_s^d}{R_s^\ddagger}\right)\right] \tag{5}$$

where n_c is the peak capacity prior to deconvolution. Eqn. 5 gives the enhancement factor in the number of peaks that are ideally obtained by deconvolving the chromatogram as viewed by the ratio p_d/p. As can be seen from eqn. 5, this enhancement becomes greater as the ratio m/n_c increases. This ratio, m/n_c, also known as the saturation factor, α, is a measure of how crowded a chromatogram is; the efficacy of deconvolution is thus seen to increase for more crowded chromatograms. An example of the value of deconvolution is suggested by considering a chromatogram with $\alpha = 1.00$ and $R_s^d = 0.35$. We choose $R_s^\ddagger$ here to be equal to 0.5 on a purely empirical basis. Under these conditions, *ca.* 1.35 times more peaks, or an increase of 35% in the number of observable peaks, will be obtained. This suggests that a distinct beneficial gain could be obtained by considering deconvolution in the analysis of chromatograms with crowded peaks. Evaluation of eqn. 5 with a variety of α values is shown in Fig. 2. As can be seen, medium α chromatograms are predicted to benefit

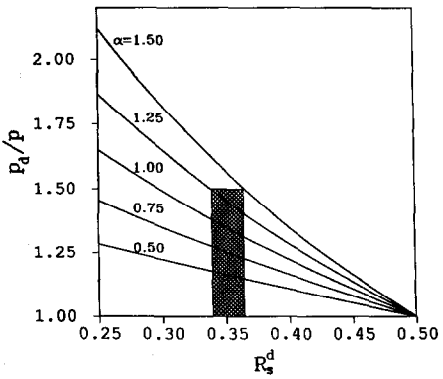

Fig. 2. The ratio p/p_d as a function of R_s^d at various saturation values, α, from eqn. 5. The critical resolution, $R_s^\ddagger$, is taken to be 0.5. The shaded bar is explained under Results and Discussion.

from deconvolution, although to a lesser extent than large α chromatograms. For relatively uncrowded chromatograms (small α), where the number of components is low or the peak capacity is high, deconvolution is known to allow better peak quantification by defusing overlapped peaks; this role of chromatographic deconvolution has been reported many times (*e.g.*, [15–20]). In addition, deconvolution has long been used in size-exclusion chromatography (SEC) where the inherently low resolution of the technique is commonly augmented by resolution enhancement [21]. In this paper, our aim is to examine super-resolution, a term commonly used in signal processing to denote the deconvolution or estimation of signals below the physically imposed broadening limit. For chromatography, the definition is broad but has recently been clarified to include the resolution of peaks separated by less than 2σ in time or alternatively with R_s less than 0.5 [22].

Problem formulation

Consider a series of Gaussian peaks composing a chromatogram of m components eluting along the time axis t so that

$$F(t) = \sum_{j=1}^{m} \frac{B_j}{\sqrt{2\pi}\sigma_j} \exp\left[\frac{-(t - t'_j)^2}{2\sigma_j^2}\right] \tag{6}$$

where $F(t)$ is the observed detector response at time t and B_j is the detector response factor of each component j with retention time t'_j and variance σ_j^2. The variance is expressed here in units of (time)^2 and the detector response factor is normalized so that $\int_0^\infty F(t)\,dt = B_j$ for the elution of the pure jth component. It is assumed here that the detector response, $F(t)$, is the linear combination of amplitudes from each component, which is an excellent assumption for detectors such as the conventional UV detector used in HPLC. The peak variance, σ_j^2, is the result of ideal intra- and extra-column broadening processes. Initially, the sample band is introduced into the column as a narrow plug which very nearly approaches a delta function, $\delta(t)$, compared with the width of a zone at the detector. Using previously developed theory [23,24], elution can be modeled as a convolution process. For ideal one-component elution with Gaussian broadening,

$$F(t) = \int_{-\infty}^{\infty} G(t - t')\delta(t)dt' \tag{7}$$

where G is the broadening operator, normalized so that eqn. 7 is independent of the measurement space, and t' is an auxiliary variable. A shortened version of eqn. 7 is often used [25,26]:

$$F(t) = G(t) * \delta(t) \tag{8}$$

where the $*$ operator denotes convolution. Using the well known commutative property of convolution [25,26]:

$$\int_{-\infty}^{\infty} G(t-t')W(t')\mathrm{d}t' = \int_{-\infty}^{\infty} G(t')W(t-t')\mathrm{d}t' \tag{9}$$

where W is some function operated on by the broadening operator, eqn. 7 can now be expressed as

$$F(t) = \int_{-\infty}^{\infty} G(t')\delta(t-t')\mathrm{d}t' \tag{10}$$

Physically, under the most ideal circumstances, this representation allows G to be a pure band broadening operator and the delta function now represents the delay time invoked by retention of the solute. Hence t' is identified as the retention time, t_r. If a mathematical inversion process were known so that given $F(t)$ and the functional form of G, t' could be determined with exact accuracy, then chromatography could be performed where all peaks in a chromatogram, regardless of the degree of fusion, could be recognized and quantified. Owing to the ill-conditioned nature of the problem, this is not possible [22]. If, in principle, eqn. 10 could be exactly solved then R_s^d would be zero because this situation corresponds to a zero critical resolution and eqn. 5 would ideally give $p_a/p = \exp(m/n_c)$, the maximum amount of resolution enhancement. This is not currently possible, and implies, in the discrete equation analogue of eqn. 10, an infinite sampling rate because the $\delta(t-t')$ term would have to lie exactly at the retention time of the peak. The presence of noise also foils this scheme. Using approximations instead of exact solutions for eqn. 10 can yield useful information if the $\delta(t-t')$ term is replaced with the function $W(t-t')$, which will now be defined as a function resembling a δ function but with finite width. Utilizing approximate solutions of eqn. 10 to solve for $W(t-t')$, as will be shown in the Results section, can give very useful information.

Numerical methodology

Although a host of numerical methods have been utilized for chromatographic and spectroscopic deconvolution, including those based on Fourier transform [15,16], least-squares solution [20], Kalman filters [27], singular value decomposition [28,29], maximum entropy methods [29], constrained iterative relaxation methods [19,30–34] and mixtures of these methods [29], a common theme that has emerged is that each method has its own advantages and disadvantages. For instance, Fourier transform methods, although fast and simple to implement, offer for most studies only a small improvement in resolution. Alternatively, the maximum entropy method (MEM) and the constrained iterative relaxation method (CIRM) offer the possibility of high deconvolution power while maintaining a reasonable baseline and peak integrity at the expense of increased computer time and increased implementation complexity. For the CIRM, high-resolution deconvolution is accomplished because the constraints of positivity and maximum value allow frequency extrapolation to occur [34], *i.e.*, the power spectral density function present in the native signal can be moved into a higher frequency domain without spurious peak generation or distortion under the best conditions. The degree to which frequency extrapolation can take place ultimately limits the minimum value that σ_d^2, the variance of the deconvolved peak, can obtain under reasonable fidelity. Fourier transform methods do not inherently allow

frequency extrapolation to occur unless constraints, such as non-negativity of solution, are combined with the method [18,35].

We now focus on the CIRM because it will be used in this paper for deconvolving synthetic chromatograms. This method is performed on a chromatogram by applying the discrete point formula:

$$W^{(k+1)}(t_i) = W^{(k)}(t_i) + r\{W^{(k)}(t_i)\}\{F(t_i) - [G * W^{(k)}]|_{t_i}\} \tag{11}$$

where the superscript denotes the iteration number, the subscript denotes the position in each vector corresponding to a unique time value and the term $r\{W^{(k)}(t_i)\}$ is the relaxation function [30]. Prior to applying this algorithm, $F(t_i)$ is scaled to between 0 and 1 and the initial values of $W^{(0)}(t_i)$ are set equal to $F(t_i)$. The functional form of the relaxation function used here is

$$r\{W^{(k)}(t_i)\} = r_0 W^{(k)}(t_i)[1 - W^{(k)}(t_i)] \tag{12}$$

where r_0 is a constant, usually equal to 2 [30]. It is easy to see how this algorithm works. The difference vector formed by $F(t_i) - [G * W^{(k)}]|_{t_i}$ is used to iteratively form each new estimate of $W^{(k+1)}(t_i)$, hence when the difference vector is small, as convergence is reached, $W^{(k+1)}(t_i) \approx W^{(k)}(t_i)$. Furthermore, the relaxation function maintains the positivity and finite constraints on $W^{(k+1)}$. Further information on this method is described by Jansson [30].

It is well known that the convolution operation can be implemented with fast Fourier transform (FFT) techniques [26]. A criticism of all numerically intensive algorithms based on convolution is that unless FFT methods are utilized for convolution, long execution times will result, which would prohibit the use of convolution on laboratory computers. This is because the computational complexity of discrete convolution, which is of the order M^2 floating-point operations (where M is the number of points in the chromatogram), grows fast with increasing M, as compared with the computational complexity of the FFT, which is of the order $M\log M$. For the utmost flexibility, however, FFT methods cannot convolve $W(t)$ with a broadening function that varies with time, because the FFT is a time-invariant operator [26].

For the case where time-varying broadening occurs across the chromatogram, for instance in isocratic LC or isothermal GC, the G term in eqn. 11 is actually composed of an M by M matrix [25] with the row number i corresponding to a unique time t in the fractogram and the column number j corresponding to a unique retention of a peak (the t' function of a peak). Each matrix entry therefore represents the broadening amplitude which occurs at time t_i when the retention time is t_j. Hence the convolution of G with W is simply a matrix multiplication [25] of the matrix G, of size M by M, with the vector W, of length M. Because these broadening response sequences contained in the matrix G are mostly zero, over the full duration of the response, there is no need to store these zeroes. In practice, an index table for the start of each column's non-sparse broadening function is utilized and the dense part of G is compressed into a banded vector. A special multiplication algorithm then forms the $G * W$ convolution. The speed and storage savings with this scheme allow relatively fast operation with minimum memory requirements. For example, a 20-min chromatogram sampled

at two points per second with a constant Gaussian response of $\sigma = 6$ s would normally require $5.76 \cdot 10^6$ entries in memory for the G matrix. If only $\pm 4\sigma$ of the Gaussian broadening curve is stored in memory, then the compressed G storage requires only 230 400 entries, a saving of a factor of 25 as compared with the full storage scheme. The speed increase with G compressed is also a factor of 25 faster than convolution with the full G matrix of size M by M. This makes the method practical to use for chromatography with fast workstations exceeding 10^6 floating point operations per second.

Implementation

The computer programs used here are written in FORTRAN-77 and run on a variety of UNIX and VMS operating system computers. Digital filtering is performed initially on $F(t_i)$ and on consecutive iterations to the convolution pair in eqn. 11 with a Kaiser filter [36] set to roll off at between 0.35 and 0.5 of the Nyquist sampling frequency. This filter has smooth roll-off characteristics and avoids the oscillatory behavior that is known to occur with the Savitsky–Golay filters in the frequency domain [37]. If filtering is not applied on every iteration, peak shapes become distorted with rectangular character as σ_d becomes small; because the filtering tends to round the rectangular edges of peaks through high-frequency attenuation, this may be only a cosmetic situation. The convergence rate appears to be little affected by the choice of cut-off frequency once filtering is performed on the original signal because the band-limiting criterion needed for convergence [34] is established. However, stability and the minimum σ_d obtainable from this method are no doubt balanced by the proper choice of the roll-off frequency and the sharpness of the filter roll off. In this regard, the digital filter used in the CIRM has to be properly chosen; in this study we chose filter conditions where stability was the primary concern.

Typically, 80 iterations of the CIRM are used for deconvolution; however, in some instances 160 iterations are used because of slow convergence. The root mean square (r.m.s.) error, computed from the difference vector, is monitored at each iteration and a smooth exponential-like decrease in the r.m.s. error is noted to occur as a function of iteration number. Typically, the discrete point density is maintained to be *ca.* 36 points over the peak (extending $\pm 4\sigma$ from the peak mean). This oversampling [15,16] promotes high accuracy in the convolution operation; preliminary experiments demonstrate reasonable peak fidelity down to 20 points over the native peak when $\sigma/\sigma_d \approx 5$.

RESULTS AND DISCUSSION

Maximum resolution deconvolution

To probe the extent of deconvolution that is possible with the CIRM, synthetic chromatograms composed of two peaks are generated with specific peak spacings and peak-height ratios and subsequently deconvolved. In the results given here only noiseless Gaussian peaks are considered; the effects of skewed peaks and noise upon deconvolution will be examined in another paper. These results are oriented towards revealing peaks that are fused in the native chromatogram, hence accuracy in peak areas is not measured here.

The result of these experiments is shown in Fig. 3, where R_s^d is given as a function

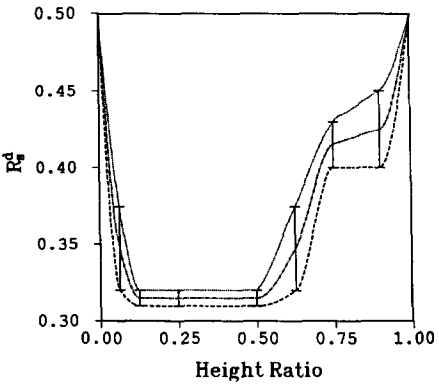

Fig. 3. Limits of deconvolution on two-component Gaussian peaks as embodied in the deconvolved peak critical resolution, R_s^d, as a function of the peak-height ratio, h_r. The bars and spline fits are explained in the text.

of the peak-height ratio, h_r. The region at the top of the bars is where the two-component peaks are easily deconvolved in 80 iterations and the peaks are resolved to within 20% of the baseline, as measured from the smallest peak height. The region at the bottom of the bars represents the threshold of deconvolution where peaks are no longer able to be separated numerically, with respect to the visual presence of two peaks. Between the bars two peaks are discernible but become increasingly difficult to differentiate as two peaks as the distance between the peaks is decreased. Most deconvolution operations in the region between the two bars requires 160 iterations for peak resolution and convergence to an r.m.s. error of $<2 \cdot 10^{-3}$.

The three broken lines shown in Fig. 3 are calculated from spline fits [38] of the top, bottom and mean of these limit bars. The function $R_s^d(h_r)$ is extrapolated at both low and high h_r to 0.5. For a doublet of equal peak height it becomes impossible normally to distinguish two discrete peaks below $R_s = 0.5$ [39]; however, the CIRM cannot deconvolve doublet peaks of equal height at this resolution; this will be discussed below. At low h_r, the method appears to approach asymptotically the resolvability limit of $R_s^d(0) = 0.5$, although this limit is questionable because the peak-height ratio of zero has no meaning and the function $R_s^d(h_r)$ is singular at $h_r = 0$. The curve shapes outlined by the spline interpolation show that the CIRM is, on average, more effective at small than large h_r. The non-symmetrical nature of Fig. 3 is partly explained by the shape of the function describing the resolution of native peaks as a function of the height ratio, which will be discussed below. In addition, the non-symmetrical response may have its origin partly in the non-linear operation of the deconvolution method.

The spline coefficients obtained from fitting are conveniently used to compute estimates of the average critical resolution of deconvolution, $\overline{R_s^d}$, over the range $0 < h_r \leqslant 1$ by a weighted integration to find the function average value:

$$\overline{R_s^d} = \int_{0+}^{1} w(h_r) R_s^d(h_r) \mathrm{d}h_r \tag{13}$$

where the lower integration limit is set to be infinitesimally larger than zero because of the singularity mentioned previously. The term $w(h_r)$ is a weighting factor used to compensate for the non-uniform probability of height ratios sampled from exponentially distributed random variables; recent experiments have suggested that the peak area follows an exponential probability density function [7] for complex samples. Additional assumptions governing this exponential density under constant zone width and with random detector response factors have recently been discussed [5,7–10].

For most chromatographic applications, the peak height ratio, h_r, is usually defined for two peak heights, h_1 and h_2, to be $h_r = \min\{h_1,h_2\}/\max\{h_1,h_2\}$ and this convention is used throughout this paper. The normalized weighting function for the ratio of exponential density random variables has been given by Feller [40] for the case where $0 < h_r \leqslant \infty$, *i.e.*, the numerator and denominator of the ratio can take on all values except zero in the denominator. The normalized density function which we shall use for weighting is simply twice the function given by Feller (this has been verified by computer simulation):

$$w(h_r) = \frac{2}{(1 + h_r)^2} \tag{14}$$

For peaks where the height ratio density function is uniform, the weighting function, $w(h_r)$, is unity. In both the exponential and uniform density cases the values of $w(h_r)$ given here apply only to random peak heights with no serial correlation.

Evaluation of eqn. 13 for the curves displayed in Fig. 3 is given in Table I. The results in Table I and Fig. 3 indicate that the CIRM is capable of deconvolving, under zero noise and in the absence of tailing, to super-resolution ($R_s < 0.5$) with the exception of peaks with $h_r = 1$ and spaced less than $\Delta t_r = 2\sigma$. For the case of multiple peaks (triplets, quadruplets, etc.) with adjacent $h_r = 1$ and $R_s \leqslant 0.5$, where the peak shape is essentially flat at the top of the peak aggregate, the CIRM appears to be incapable of resolving any of the aggregate. This probably results because the relaxation function given in eqn. 12 is zero when $W^{(k)}(t_i) = 1$. For this extreme case, forcing the relaxation function to take on non-zero values where $W^{(k)}(t_i) = 1$ does not help because the iterative process guides the peak top to the region where the relaxation function assumes a zero value. There appears to be no easy modification to the CIRM that will maintain the top constraint and prevent this condition from occurring, short of adding noise to prevent the flat peak top condition.

Statistical chromatogram deconvolution

In an attempt to evaluate the CIRM numerical methodology on complex

TABLE I

EVALUATION OF $\overline{R_s^d}$ FOR CIRM FROM TWO-COMPONENT PEAKS

Limit	Uniform h_r weighting	Exponential h_r weighting
Low limit	0.348	0.339
High limit	0.375	0.364
Mean limit	0.362	0.352

chromatograms, synthetic chromatograms with 20–72 peaks at different α values are produced with subsequent deconvolution by the CIRM. The retention times are generated with a Poisson distribution, as has commonly been used in previous studies of statistical peak overlap [2–13], and with an exponential density of peak heights [5,7–10]. Constant peak width is used in these calculations; however as stated previously, constant peak width is not a limitation of the CIRM or the convolution algorithm used here. Because of the statistical nature of these numerical experiments, it is necessary to perform a number of calculations at each stated saturation ratio, α; random number seeds are varied within these sets to produce estimates of the ratio $\overline{(p_d/p)}$. To accommodate this requirement, five synthetic chromatograms are calculated at each α. Peak counting is performed by visible inspection of both native and deconvolved synthetic chromatograms. The presence of peak maxima and shoulders is utilized in determining the number of peaks that are present in the chromatograms.

Four cases of deconvolution at different α values are given in Figs. 4–7 using the CIRM. As can be seen, deconvolution with constant σ and in the absence of noise and tailing renders many more peaks visible than in the fused peak native chromatograms shown. As shown by the retention time markers in Figs. 4–7, the CIRM cannot remove peak fusion when retention times are very close, as explained previously. Figs. 4–7, however, clearly demonstrate that deconvolution at the super-resolution level has the potential to be an effective aid in normal chromatographic analysis.

The increase in the number of visible peaks after deconvolution, as measured by $\overline{(p_d/p)}$, the average ratio of visible deconvolved peaks to visible native peaks, is given in Table II as a function of the saturation ratio, α. As can be seen, there is a gain in the number of detectable components, with the largest increase coming from large α experiments, as predicted from eqn. 5. Table II also indicates that there is a dispersion

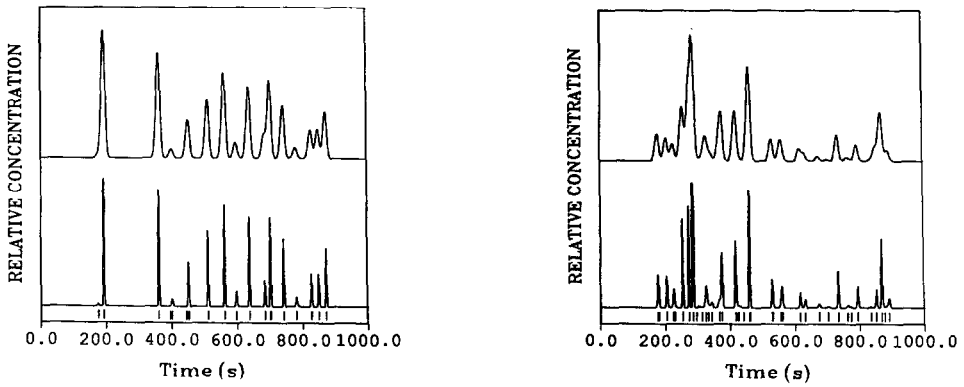

Fig. 4. Statistical chromatogram before (top) and after (bottom) deconvolution. Conditions: $\alpha = 0.513$ (with $R_s^{\dagger} = 0.8$ in n_c), $m = 20$, $\sigma = 6$ s, 80 iterations, 15 peaks detected in the native chromatogram and 16 peaks detected in the deconvolved chromatogram. Retention times are indicated by the bars below the bottom chromatogram.

Fig. 5. Statistical chromatogram before (top) and after (bottom) deconvolution. Conditions: $\alpha = 1.00$ (with $R_s^{\dagger} = 0.8$ in n_c), $m = 39$, $\sigma = 6$ s, 80 iterations, 21 peaks detected in the native chromatogram and 28 peaks detected in the deconvolved chromatogram. Retention times are indicated by the bars below the bottom chromatogram.

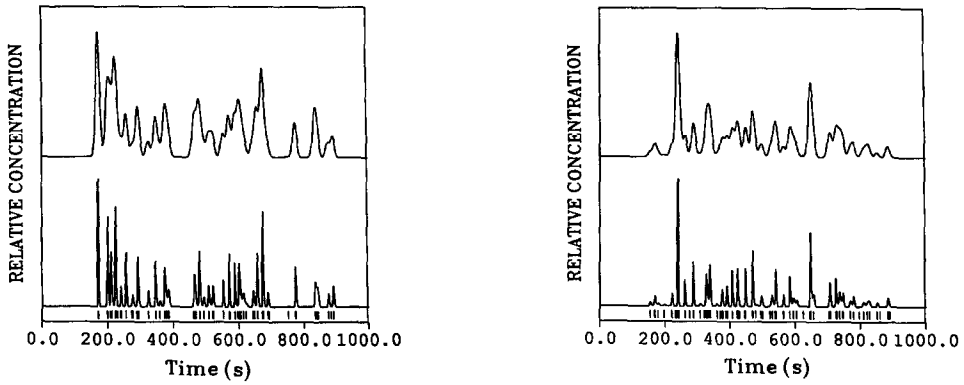

Fig. 6. Statistical chromatogram before (top) and after (bottom) deconvolution. Conditions: $\alpha = 1.49$ (with $R_s^{\ddagger} = 0.8$ in n_c), $m = 60$, $\sigma = 6$ s, 160 iterations, 22 peaks detected in the native chromatogram and 33 peaks detected in the deconvolved chromatogram. Retention times are indicated by the bars below the bottom chromatogram.

Fig. 7. Statistical chromatogram before (top) and after (bottom) deconvolution. Conditions: $\alpha = 1.85$ (with $R_s^{\ddagger} = 0.8$ in n_c), $m = 72$, $\sigma = 6$ s, 160 iterations, 25 peaks detected in the native chromatogram and 38 peaks detected in the deconvolved chromatogram. Retention times are indicated by the bars below the bottom chromatogram.

in $\overline{(p_d/p)}$, as expected, although it is not large as viewed from the relative standard deviation column. In an attempt to relate $\overline{(p_d/p)}$ from eqn. 5 to $\overline{R_s^d}$ from Table I, two different $R_s^{\ddagger}$ values are used in eqn. 5; these results are also given in Table II. It is especially important to note that $R_s^{\ddagger}$ appears in eqn. 5 explicitly as part of the ratio of critical resolutions ($R_s^d/R_s^{\ddagger}$); however, $R_s^{\ddagger}$ also appears implicitly in the n_c term in eqn. 5 through eqns. 3 and 4. For reasons that are not clear, good consistency is obtained when $R_s^{\ddagger}$, implicitly carried in n_c, is chosen to be 0.8 (consistent with [9]) and $R_s^{\ddagger}$ in the

TABLE II

EVALUATION OF $\overline{R_s^d}$ FROM STATISTICAL CHROMATOGRAMS

α^a	$\overline{(p_d/p)}$	%R.S.D.[b] in (p_d/p)	$\overline{R_s^{dc}}$	Deviation from $R_s^d = 0.352$ (%)	$R_s^{\ddagger}$ [d]
0.513	1.13	7.23	0.383	8.91	0.539
0.794	1.28	8.77	0.346	−1.73	0.598
1.00	1.33	12.0	0.357	1.36	0.581
1.18	1.44	10.2	0.345	−2.12	0.601
1.49	1.51	8.44	0.361	2.61	0.574
1.67	1.42	6.35	0.394	12.0	0.521
1.85	1.48	9.78	0.394	11.9	0.522

[a] Calculated with $R_s^{\ddagger} = 0.8$ in n_c term.
[b] R.S.D. = Relative standard deviation.
[c] Assuming $R_s^{\ddagger} = 0.5$ in ratio term.
[d] Assuming $R_s^d = 0.352$ in ratio term, $R_s^{\ddagger}$ applied equally to ratio and n_c in α.

ratio $R_s^d/R_s^\ddagger$ is equal to 0.5 (consistent with [3] and [11]). As is noted from Table II, deviation of R_s^d from the value 0.352 is very low for the numerical experiments where $\alpha \leqslant 1.49$. In addition, R_s^d can be accepted as 0.352 and then the value of the critical resolution, $R_s^\ddagger$, can be obtained through rearrangement of eqn. 5; these values are shown in the last column in Table II. It is noted that in this instance $R_s^\ddagger$ varies over the range $0.574 \leqslant R_s^\ddagger \leqslant 0.601$ when $0.794 \leqslant \alpha \leqslant 1.49$. From the limited data given in Table II, there does not appear to be a recognizable trend which can be used in establishing one uniquely consistent value of $R_s^\ddagger$. However, the empirical use of 0.8 for $R_s^\ddagger$ internal to n_c and $R_s^\ddagger$ of 0.5 in the ratio $R_s^d/R_s^\ddagger$ serves to allow estimation of the ratio p_d/p given R_s^d from the deconvolution methodology at a stated α. The value of 0.352, from Table I, is independent of the statistical peak overlap model; hence this gives further credibility to the assignment of $\overline{R_s^d} = 0.352$ for the CIRM and suggests an independent verification for the simple statistical model of peak overlap with $R_s^\ddagger = 0.5$ [2,3,9,11]. In Table II, the lowest α result appears to be in poor agreement with $\overline{R_s^d} = 0.352$; a possible explanation is that the number of peaks in this experiment is small; only 20 peaks are used at this α value and hence the statistical sampling is poor with only five experiments contributing to the mean value. When the statistical peak overlap model is used for the estimation of the true number of peaks in both synthetic and experimental chromatograms, it is observed [11] that the accuracy of estimation degrades at $\alpha > 0.5$. It is possible that the additional mathematics needed to describe higher saturation effects (*e.g.*, the perturbation of the peak heights due to closely spaced surrounding clusters) will be less important in the application given here because the use of ratios, on which eqn. 5 is based, may tend to cancel or minimize this additional level of theory.

The results in Table II illustrate the monotonic increase in $\overline{p_d/p}$ from statistical chromatograms for $\alpha \leqslant 1.49$. Above this value, however, there is no longer an increase in p_d/p and it is at this point that the CIRM loses the ability to quantitatively remove peak fusion. A number of numerical experiments conducted with $2.0 \leqslant \alpha \leqslant 100.0$ have been used to observe that for $\alpha \geqslant 3.0$ the ratio p_d/p is consistently approximated by the value 2.1 and deconvolution gives meaningless results in this range. For the range $1.5 \leqslant \alpha \leqslant 3.0$, p_d/p asymptotically, but not monotonically, approaches 2.1 with broader deconvolved peaks than shown in Figs. 4–7. Therefore, in the absence of noise and tailing, the CIRM is observed to be effective, in the statistical sense, up to α values less than 1.5 and with the potential visible restoration of 1.5 times the number of peaks appearing in the native chromatogram. This operating range is summarized in Fig. 2 in the shaded area.

It appears from the previous discussion that deconvolution can quantitatively restore lost peak information; however, as predicted from the statistical model of peak overlap [2–13], the general performance of separation is poor as analysis is carried out at higher α conditions. This is suggested by the data in Table III, where the quantities $\overline{p/m}$ and $\overline{p_d/m}$ are given from statistical chromatogram production and deconvolution. It is noted here that although deconvolution provides an enhancement in the ratio of the number of visible peaks after deconvolution to the true number of components, this ratio is still far below unity, the value that would be obtained if deconvolution could resolve all the peaks in the chromatogram. It is seen from Table III that there is a gain, however, in that deconvolution allows working at higher α for a constant ratio

TABLE III

EVALUATION OF $\overline{p/m}$ AND $\overline{p_d/m}$ FROM STATISTICAL CHROMATOGRAMS

α^a	m	$\overline{p/m}$	$\overline{p_d/m}$
0.513	20	0.720	0.810
0.794	31	0.587	0.742
1.00	39	0.528	0.697
1.18	46	0.461	0.661
1.49	58	0.397	0.597
1.67	65	0.372	0.529
1.85	72	0.339	0.500

[a] Calculated with $R_s^{\ddagger} = 0.8$ in n_c term.

of visible peaks to true components, as compared with the native chromatogram. For example, with $\alpha = 0.513$, $\overline{p/m} = 0.720$ for the native chromatogram. With deconvolution, $\overline{p_d/m} = 0.697$ for $\alpha = 1.00$. These results suggest that CIRM deconvolution, although not capable of restoring all the peaks to a 1:1 identity with the components, can be used to perform separations at higher column saturation values than would normally be used for high-resolution separations.

Although our attention is focused on the extent of deconvolution of complex chromatograms in this paper, it is instructive to briefly examine additional $\overline{R_s^{\ddagger}}$ values for the statistical model which can be produced from the average critical resolution equation developed in this paper. As demonstrated by El Fallah and Martin [10], synthetic chromatogram methodology may be used to estimate $\overline{R_s^{\ddagger}}$ in the context of the statistical model of peak overlap. A different approach, as used here, is to find $\overline{R_s^{\ddagger}}$ independent of the statistical model.

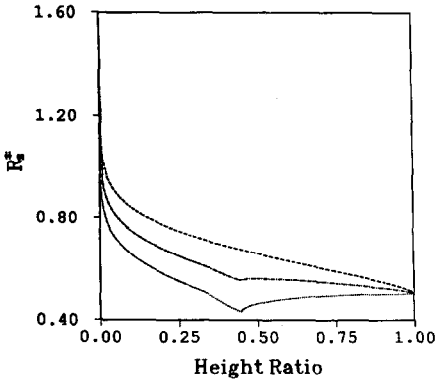

Fig. 8. Limits of resolving two-component Gaussian peaks as a function of the peak-height ratio, h_r. The top curve is the critical resolution value where valley discrimination between two peaks ceases. The bottom curve is the critical resolution value where inflection points can no longer be used to discriminate between two unique peaks. The middle curve is the average of the top and bottom curves.

In an attempt to find one uniquely consistent $\overline{R_s^\ddagger}$ value, critical resolutions for a two-component Gaussian doublet are calculated with eqns. 13 and 14 using $R_s^\ddagger(h_r)$ from two limiting systems: (1) the values of $R_s^\ddagger(h_r)$ where the valleys between fused peaks vanish as resolution is decreased and (2) the values of $R_s^\ddagger(h_r)$ where inflection points can no longer distinguish between peaks as resolution is decreased. The former case has been called the "shoulder" resolution case and the latter the "detectability" resolution case [15]. Analytical solutions for these two situations were given by Westerberg [15]. These two situations are graphically illustrated for $R_s^\ddagger$ as a function of h_r in Fig. 8 and the evaluation of these two cases using eqns. 13 and 14 is given in Table IV.

The $\overline{R_s^\ddagger}$ values for exponential weighting from Table IV are consistent with bracketing the values of $0.574 \leqslant \overline{R_s^\ddagger} \leqslant 0.601$ when $0.794 \leqslant \alpha \leqslant 1.49$ from the last column in Table II. This suggests some form of consistency in accepting the value of $\overline{R_s^\ddagger}$ of *ca.* 0.645 (the mean limit from Table IV) when eqn. 5 is formulated in a fully consistent usage of $R_s^\ddagger$. Unfortunately, this does not give insight into the proper value of $R_s^\ddagger$. Of interest, however, is the correspondence of values in Table IV with values calculated within the statistical model framework [10]. In this instance El Fallah and Martin [10] obtained values of $\overline{R_s^\ddagger} = 0.536$ for the low limit (compared with 0.558 in Table IV) and $\overline{R_s^\ddagger} = 0.796$ for the high limit (compared with 0.725 in Table IV). Although direct comparison of these numbers is not strictly valid because of differences in the definition of the measurements and because the $\overline{R_s^\ddagger}$ values estimated in Table IV only include fused doublet peaks, the correspondence is nonetheless noted. There is, however, an interesting correspondence between El Fallah and Martin's value of 0.71 ± 0.01 for the average resolution, which is suggested for use [10] in the statistical peak overlap model, and the average resolution calculated for the shoulder resolution case in Table IV of 0.725. Towards this end, further theoretical work will be needed to clarify an exactly rigorous value (if one exists) for $\overline{R_s^\ddagger}$. As was stated previously, good consistency is realized between the results in Table II and from eqn. 5 when $R_s^\ddagger$, implicitly carried in n_c, is chosen to be 0.8 and $R_s^\ddagger$ in the ratio $R_s^d/R_s^\ddagger$ is chosen as 0.5, although at this point these values must be viewed as strictly empirical.

Overview

A comparison may be made between the results given here and investigations into overlapping peak resolution with derivative methods. The derivative methods [41], tend to be noise sensitive, as is the CIRM. However, CIRM tends to be more

TABLE IV

EVALUATION OF $\overline{R_s^\ddagger}$ FOR TWO-COMPONENT PEAKS AT THE SHOULDER AND DETECTABILITY LIMITS

Limit	Uniform h_r weighting	Exponential h_r weighting
"Detectability" limit	0.529	0.558
"Shoulder" limit	0.678	0.725
Mean limit	0.605	0.645

effective in resolving severely overlapped peaks below $R_s = 0.5$. This is, of course, at the expense of computer time, where the derivative methods typically require 1 s or less of computer time and the CIRM method takes about 5 min. In addition, derivative methods require little or no information about the column response function, which is a critical matter in the use of CIRM or any other method based on convolution. If the two methods could be combined, for instance if derivative information could be used to obtain a better starting estimate of $W^0(t_i)$, the deconvolution method may be made more time efficient, requiring less iteration to reach convergence.

The response function embodied in the G matrix may be obtained experimentally by injection of individual solutes or by injection of a mixture when baseline separation is possible. In this way peak parameters can be estimated and used to form G. As stated previously, the CIRM methodology is capable of handling time-dependent broadening in the composite peak shape throughout the chromatographic elution profile. In addition, it may be possible, depending on the nature of the solutes in the test and chromatographic mixtures, to substitute similar solutes in the test mixture as long as the elution range is covered. Careful experiments must be performed, in fact, to determine whether this is feasible and to what class of substances this can be applied. In essence, this may depend on whether specific interactions are controlling the separation and whether tailing is an extra- or an intra-column effect (*e.g.*, slow desorption from the stationary phase). The reproducibility of the solvent or temperature program may also affect the accuracy of the peak response profile when gradient elution (LC) or temperature programming (GC) is used.

One of the many aspects of CIRM deconvolution not discussed in this paper is the effect that inaccurate σ values (and other peak model parameters when non-Gaussian peaks are encountered) will have on the resultant deconvolved chromatogram. Synthetic chromatograms have been produced and analyzed where the broadness is purposely reduced in the G matrix; on deconvolution no spurious peaks are recognized but there are noticeably larger σ_d values and poor overall performance. In instances where the broadening is purposely made larger than the σ value in the native singlet peak, occasional splitting into doublets has been observed, but this only happens after extensively long iteration. In this instance the rate of convergence is seen to be very slow and sometimes divergence occurs, as viewed from plot of r.m.s. error *versus* iteration number. The characterization of the broadening response of pure components in experimental chromatograms may be the accuracy-limiting step in the CIRM deconvolution process; this may also be true of any deconvolution method based on convolution.

The results given here were obtained under the limiting condition of no noise. Noise is well known to be a critical factor in limiting the ability of deconvolution algorithms to work properly [42] and this factor must be determined experimentally because of the unique nature of the noise frequency distribution in real chromatograms. Further, it has been suggested [30] that most deconvolution methods will work well on synthesized peaks but may fail when real waveforms are used. As shown in this paper, there is a finite resolving capability of the CIRM, and the purpose of this study was to find the region(s) of operation where resolution could not be restored. In addition, the Fourier method of deconvolution has been analyzed in the context of "perfect" experiments [16], where it was shown that delta function recovery was impossible and that only a small amount (compared with the CIRM) of resolution

recovery was possible. If the deconvolution methodology cannot enhance the resolution of noiseless signals, it is doubtful that high-fidelity resolution enhancement can be performed on a real noise-containing signal. There is no doubt, however, that the extent of deconvolution presented in this paper can be approached with very careful experiments, modern chromatographic equipment, and rigorous digital filtering when concentrations are adequate to permit working above the detection limit. In the context of filtering, it is known that CIRM deconvolution can be performed when the native detector signal is significantly broadened by overfiltering [34]; owing to the frequency extrapolation that occurs with CIRM, additional filtering can be deconvolved from the signal by including the filter response in the broadening operator, G.

The procedures for statistical chromatogram deconvolution presented in this paper may also be useful for evaluating other deconvolution methods in the context of chromatography. For instance, deconvolution by the maximum entropy method or the relatively new method of deconvolution by splines [43] may be evaluated by calculating $\overline{R_s^d}$, from R_s^d *versus* h_r numerical experiments, and utilizing eqn. 5 to relate α and $\overline{R_s^d}$ to p_d/p. Finally, we note that the CIRM may have application in reducing analysis time whereby the chromatographic flow velocity may be increased at the expense of losing resolution, the resolution loss being compensated by numerical deconvolution. Chromatographic experiments are in progress to determine the extent to which this is feasible.

SYMBOLS

B_j	Detector response of component j in a complex mixture
$F(t)$	Time-based concentration response of detector
$G(t), G(t,t')$	Broadening operator
h_r	Peak-height ratio
h_1, h_2	Heights of peaks 1 and 2
j	Summation index
m	Number of true components present in a mixture
M	Number of points in a digitized chromatogram
n_c	Peak capacity
n_c'	Peak capacity at unit resolution
N	Number of theoretical plates
N_{av}	Average number of theoretical plates
$\underline{p}$	Number of visible peaks in native chromatogram
p	Average number of visible peaks in native chromatograms
$\underline{p_d}$	Number of visible peaks after deconvolution
p_d	Average number of visible peaks after deconvolution
r	Relaxation function
r_0	Relaxation constant
$\underline{R_s^d}$	Critical resolution of deconvolution
$\overline{R_s^d}$	Average critical resolution of deconvolution
$\underline{R_s^\dagger}$	Critical resolution value for peak discrimitation
$\overline{R_s^\dagger}$	Average critical resolution value for peak discrimination
t, t_r	Time, retention time

t_{max}	Time at the end of chromatographic separation
t_0	Void time of a chromatogram
$w(h_r)$	Weighting factor for peak heights
$W^{(k)}(t_i)$	kth iterative estimate of the unbroadened peak function
α	Saturation factor, $\alpha = m/n_c$
$\delta(t)$	Delta function
σ,σ_j	Standard deviation of a Gaussian peak
σ_d	Standard deviation of a Gaussian peak after deconvolution

REFERENCES

1 K. A. Connors, *Anal. Chem.*, 46 (1974) 53.
2 J. M. Davis and J. C. Giddings, *Anal. Chem.*, 55 (1983) 418.
3 J. C. Giddings, J. M. Davis and M. R. Schure, in S. Ahuja (Editor), *Ultra-High Resolution Chromatography (ACS Symposium Series*, No. 250), American Chemical Society, Washington, DC, 1983, Ch. 2, pp. 9–26.
4 J. M. Davis and J. C. Giddings, *J. Chromatogr.*, 289 (1984) 277.
5 M. Martin and G. Guiochon, *Anal. Chem.*, 57 (1985) 289.
6 J. M. Davis and J. C. Giddings, *Anal. Chem.*, 57 (1985) 2168.
7 L. J. Nagels and W. L. Creten, *Anal. Chem.*, 57 (1985) 2706.
8 M. Martin, D. P. Herman and G. Guiochon, *Anal. Chem.*, 58 (1986) 2200.
9 W. L. Creten and L. J. Nagels, *Anal. Chem.*, 59 (1987) 822.
10 M. Z. El Fallah and M. Martin, *Chromatographia*, 24 (1987) 115.
11 S. L. Delinger and J. M. Davis, *Anal. Chem.*, 62 (1990) 436.
12 A. Felinger, L. Pasti and F. Dondi, *Anal. Chem.*, 62 (1990) 1846.
13 A. Felinger, L. Pasti, P. Reschiglian and F. Dondi, *Anal. Chem.*, 62 (1990) 1854.
14 J. C. Giddings, *Anal. Chem.*, 39 (1967) 1027.
15 A. W. Westerberg, *Anal. Chem.*, 41 (1969) 1771.
16 D. W. Kirmse and A. W. Westerberg, *Anal. Chem.*, 43 (1971) 1035.
17 M. Rosenbaum, V. Hanch and R. Komers, *J. Chromatogr.*, 294 (1984) 31.
18 G. H. Rayborn, G. M. Wood, B. T. Upchurch and S. J. Howard, *Am. Lab.*, Oct. (1986) 56.
19 P. B. Crilly, *J. Chemometr.*, 1 (1987) 175.
20 A. H. Anderson, T. C. Gibb and A. B. Littlewood, *J. Chromatogr. Sci.*, 8 (1970) 640.
21 S. T. Balke, *Quantitative Column Liquid Chromatography—A Survey of Chemometric Methods*, Elsevier, New York, 1984.
22 J. Callis, in S. Ahuja (Editor), *Ultra-High Resolution Chromatography (ACS Symposium Series*, No. 250), American Chemical Society, Washington, DC, 1983, Ch. 12, pp. 171–198.
23 J. C. Sternberg, *Adv. Chromatogr.*, 2 (1966) 205–270.
24 J. C. Giddings, *Dynamics of Chromatography, Part I, Principles and Theory*, Marcel Dekker, New York, 1965.
25 R. N. Bracewell, *The Fourier Transform and its Applications*, McGraw-Hill, New York, 2nd ed., 1978.
26 A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
27 D. Alba and G. R. Meira, *J. Liq. Chromatogr.*, 7 (1984) 2833.
28 E. M. Rosen and T. Provder, *Sep. Sci.*, 5 (1970) 485.
29 K. Minami, S. Kawata and S. Minami, *Appl. Opt.*, 24 (1985) 162.
30 P. A. Jansson, *Deconvolution with Applications in Spectroscopy*, Academic Press, New York, 1984.
31 M. R. Schure, B. N. Barman and J. C. Giddings, *Anal. Chem.*, 61 (1989) 2735.
32 K. H. Michaelian and W. I. Friesen, *Appl. Spectrosc.*, 42 (1988) 1538.
33 P. B. Crilly, *J. Chemometr.*, 1 (1987) 79.
34 R. W. Schafer, R. M. Mersereau and M. A. Richards, *Proc. IEEE*, 69 (1981) 432.
35 S. J. Howard, *J. Opt. Soc. Am.*, 71 (1981) 819.
36 P. P. Vaidyanathan, in D. F. Elliot (Editor), *Handbook of Digital Signal Processing, Engineering Applications*, Academic Press, New York, 1987, Ch. 2, pp. 55–172.

37 K. R. Betty and G. Horlick, *Anal. Chem.*, 49 (1977) 351.
38 H. Akima, *J. Assoc. Comput. Mach.*, 17 (1970) 589.
39 L. R. Snyder, *J. Chromatogr. Sci.*, 10 (1972) 200.
40 W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. II, Wiley, New York, 2nd ed., 1971.
41 E. Grushka and D. Israeli, *Anal. Chem.*, 62 (1990) 717.
42 W. E. Blass and G. W. Halsey, *Deconvolution of Absorption Spectra*, Academic Press, New York, 1981.
43 A. G. B. M. Saase, H. Wormeester and A. van Silfhout, *J. Vac. Sci. Technol.*, A7 (1989) 1535.