Journal of Chromatography, 591 (1992) 19–29
Elsevier Science Publishers B.V., Amsterdam

CHROM. 23 763

Improved algorithm for resolution of overlapped asymmetric chromatographic peaks

Dae Young Youn, Sun Jin Yun and Kyung-Hoon Jung*

Centre for Molecular Science and Department of Chemistry, Korea Advanced Institute of Science and Technology, P.O. Box 150 Chongyangni, Seoul 130-650 (South Korea)

(First received May 14th, 1991; revised manuscript received September 23rd, 1991)

ABSTRACT

A simple and improved procedure to resolve an overlapped asymmetric chromatogram into its component peaks is proposed. The overlapped asymmetric peak profile was assumed to be a convolution of its component peaks, which were characterized by an exponentially modified Gaussian, and further simplified by the use of its derivative chromatogram. A new technique is suggested for initial guessing of peak parameters. The simulation study showed that peak parameters were able to be recovered within 5% deviations when the reduced resolution (*RR*), the ratio of the resolution to the critical resolution, was larger than 1.0 and were sufficient for recovery process. For RR values from 1.0 to 0.6, however, the recovery was not so efficient with the initial guessed values alone and was achieved with a non-linear least-squares routine within 3% deviations in most instances utilizing these values as initial guesses of iterations. The lower limit of *RR* for this technique was found to be 0.6. The validity of this algorithm for recovered parameters was confirmed by comparison with experimental observations and its recovery ability was found to show no more than a 1.6% deviation from true values and a 2.2% standard deviation throughout the study.

INTRODUCTION

There is growing interest in the deconvolution of an overlapped peak into its components for quantitative analysis of chromatograms. For this purpose, numerical treatments of chromatograms are commonly used in conjunction with optimization of the experimental conditions. These include geometrical methods, principal component analysis and least-squares methods.

Geometrical methods [1–4] such as perpendicular drop at the valley, tangent skimming and triangulation have been the most widely used in commercial integration systems. These methods suffer, however, from relatively low resolution and often give progressively lower resolutions for peaks with a high degree of overlapping or tailing [4]. Considerable improvement has been obtained by using more sophisticated numerical resolution techniques such as principal component and factor analysis [5] in multi-channel detector systems, *e.g.*, gas chromatography–mass spectrometry, gas chromatography–Fourier transform IR spectrometry and high-performance liquid chromatography–UV spectrophotometry. In spite of considerable improvements, these techniques are impractical for single-dimension detector systems. Least-squares methods, *i.e.*, non-linear [6–10] and linear methods [11], on the other hand, assume a fitting function to describe a real chromatogram in terms of peak shape parameters, *i.e.*, area, retention time, broadness and skewness. The iterative method, then, attempts to improve upon an initial set of peak-shape parameters by direct minimization in error space. Therefore, the success of these methods is strongly dependent on the model function, how well it represents real chromatograms, and the method for initial parameterization.

An ideal individual chromatographic peak may well be characterized by a Gaussian function provided that there are no instrumental distortions. In practice, however, the output profile has been better described by a skewed Gaussian form [12] owing to the non-equilibrium mass transfer [13], non-unifor-

0021-9673/92/$05.00

mity of solute–stationary phase interactions [14], dead volume [15], non-homogeneity of tube connection [16] and time lag of the detector response [17–19]. Some modelling techniques to describe these asymmetric distortions have been reported by many workers, including bi-Gaussian [20,21], Poisson [20,21], Gram–Charlier [22] and the exponentially modified Gaussian peak model (EMG) [4,9,20–26]. The EMG model has been found to represent a real chromatogram reasonably well through extensive tests both experimentally and theoretically in the authors' laboratory [9,26].

In this paper we present an improved method to describe an overlapped asymmetric peak profile as a convolution of its component peaks, described by the EMG model, and simplified further by use of its derivative as an extension of our previous work on the characterization of peak parameters of a skewed Gaussian chromatogram [26]. We also propose a simplified and rapid technique to extract initial peak parameters for practical implementation on a microcomputer.

COMPUTATION

Fitting function

The EMG for a single peak can be expressed as the convolution of a normal Gaussian with an exponential decay function:

$$h(t) = \frac{A}{\tau\sigma\,(2\,\pi)^{1/2}} \cdot \int_0^\infty \exp\left[-\frac{(t - t_G - t')^2}{2\,\sigma^2}\right] \exp\left(-\frac{t'}{\tau}\right) \mathrm{d}t' \qquad (1)$$

where A, σ, t_G and τ represent the peak area, the standard deviation of a Gaussian component, the retention time of the Gaussian component and the time constant of the exponential modifier, respectively. Eqn. 1 is then resolved into a Gaussian and a tailing component by differentiating and applying Leibniz's theorem:

$$h(t) = g(t) - \tau h'(t) \qquad (2)$$

and

$$g(t) = \frac{A}{\sigma\,(2\,\pi)^{1/2}} \exp\left[-\frac{(t - t_G)^2}{2\,\sigma^2}\right]$$

where $g(t)$ and $h'(t)$ are the Gaussian and the derivative peak heights, respectively.

An overlapped peak profile $H(t)$ of an n-co-eluted component system at a given time t can be expressed as a sum of component peaks, $h_i(t)$. Now, as A, σ, t_G and τ are different from component to component, these values also have to be included as variables in a multi-component system:

$$H(t) = \sum_{i=1}^{n} [h_i(t;A_i,t_{G_i},\sigma_i,\tau_i)] \qquad (3)$$

where the subscript i denotes the ith component. Although eqn. 3 is a general function which accommodates an n-component overlapped system, the deconvolution procedure requires lengthy iterative calculations if it is used as a fitting function. This difficulty can be overcome in most real systems that we encounter by simplifying the equation further by assuming a constant τ value for all component peaks.

The tailing parameter for overlapped pairs, τ, may be expressed as a sum of the extra, τ_e, and the intrinsic, τ_{in}, column effects. The intrinsic column effect can then be broken down further into the non-equilibrium mass transfer, τ_m, and the non-uniform solute–stationary phase interaction terms, τ_{int}:

$$\tau = \tau_e + \tau_{in} = \tau_e + (\tau_m + \tau_{int}) \qquad (4)$$

In eqn. 4, as the first term is due to the instrumental geometry and does not vary from component to component, it should be same for all overlapped component peaks for a given instrumental condition. The second term is caused by the non-uniform mass transfer in the linear velocity of mobile phase components. The term will be the same for all component peaks owing to equal peak broadening under given column conditions. On the other hand, as the third term is governed by the retention mechanism of solute–stationary phase interactions, the term may be studied for two different cases of elution patterns depending on either a linear chromatogram caused by physically and chemically similar pairs or a non-linear chromatogram due to the different components. As the first and second terms in eqn. 4 are dominant over the third term in a linear chromatogram, the total τ can be considered as a constant for the convolution. However, as the third term in the case of a non-linear chromatogram, *e.g.*, with mass overloading of a component [14], is predominant over the sum of the first and second terms, the total τ of overlapping pairs is no longer a constant.

Eqn. 3 for a linear chromatogram is then reduced further to eqn. 5 as a sum of a pure Gaussian and a product of a tailing parameter τ and derivative components of the overlapped peak:

$$H(t) = \sum_{i=1}^{n} [g_i(t;A_i,t_{G_i},\sigma_i)] - \tau H'(t) \quad (5)$$

where

$$H'(t) = \sum_{i=1}^{n} [h_i'(t;A_i,t_{G_i},\sigma_i)]$$

The derivative component, $H'(t)$, can be obtained by electronic [27] or numerical differentiation [28,29]. The distorted asymmetric overlapped peak can then be described by an n Gaussian and a tailing parameter and their parameters can easily be extracted utilizing a non-linear least-squares algorithm with fast convergence. In this study the Levenberg–Marguardt algorithm [30,31] was used for these purposes and the validity of the procedure was confirmed by checking against experimental observations.

Initial parameterization

If a reasonable estimate of the peak parameters of a component peak in the overlapped profile is attainable, the remaining portion can easily be obtained by subtracting the estimated peak contribution from the overall profile.

This procedure was performed in conjunction with our previously proposed chromatographic peak parameter characterization technique for estimation of the component peak [26]. In brief, EMG parameters for a single peak were extracted by use of four or five points of the normal and derivative peak heights by solving a cubic or quartic equation of τ. These require one to choose a starting component peak, either the early or later eluted component, and a time pocket, characterized by four or five equally spaced points depending on a cubic or quartic equation. For these purposes, we chose the later eluted component as a starting peak and its tailing portion for a time pocket[a] and then subtracted the parameters from the whole peak to obtain the parameters for the remaining portion. The peak parameters obtained by solving the equation were normally very sensitive to both noise and the position of the time pocket. The different five-point time average method [26], was used to reduce the noise effect and the partial root mean square error method (PRMSE), a root mean square error for each candidate of the time pocket, to obtain the best position of the time pocket. The initial parameterization procedure was sufficient to recover whole peak parameters when a valley was present. Whereas a valley could not be seen owing to a higher degree of overlapping, the recovered peak parameters showed larger errors and had to be refined with non-linear regression using these values as initial guesses for iteration.

Data acquisition

Simulation and data acquisition of experimental observations were performed using a PC. The computation procedures are illustrated stepwise in Fig. 1.

For the simulation study, the noise-added overlapped peak, $O(t)$, was obtained from a noise-free peak and by adding normal noise to each corresponding intensity point[b]. The normal noise was generated by multiplying 1.0% of the maximum peak height by a normal random number in the range -0.5 to $+0.5$, a possible worst value in a commercial gas chromatograph, using the subroutine GASDEV, Numerical Recipes [33]. The experimental chromatogram was obtained using a home-made GC and an HP 5880A equipped with flame ionization detectors.

[a] For deconvolution of an overlapped peak, the best accuracy and resolution are attainable by performing the procedure at the least overlapped portion of the peak. These portions of overlapped pairs are the head part of early-eluted and the tail part of later-eluted component peaks. As the peak shape parameters are heavily dependent on the tail part compared with the head part of an asymmetric peak, the late-eluted component peak was chosen as the starting component peak and the tail part of the peak as a time pocket to extract the peak shape parameters in this study (see also ref. 32).

[b] The noise added overlapped peak, $O(t_i)$, is given by

$$O(t_i) = H(t_i) + \mathrm{Max}[H(t)]/100NRN_i$$

(i = 1 to total data points), where $O(t_i)$ and $H(t_i)$ are ith noise added and noise-free intensity points, respectively, and NRN_i are the normal random numbers in the range $-0.5 \leqslant NRN_i \leqslant +0.5$.

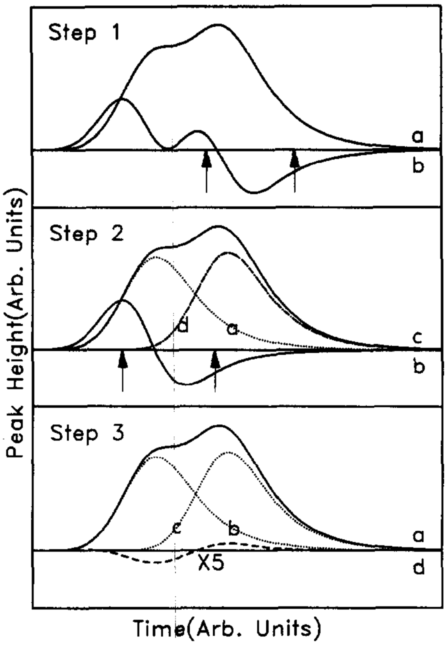

Fig. 1. Graphical representation of each step in the initial guessing routine. Step 1: the later-eluted peak parameters are extracted from the optimized time pocket of the tailing part of the overlapped peak. Solid lines represent the (a) normal and (b) derivative peaks. Arrows indicate the optimized time pocket for the later-eluted peak. Step 2: the early-eluted peak parameters are extracted from the remainder of Step 1, obtained by subtracting calculated peak contributions from the original peak. The subtracted normal, derivative, original and calculated later-eluted peaks are shown as dotted (a) and (b), solid (c), and dashed–dotted (d) lines, respectively. Arrows indicate the optimized time pocket of the early-eluted peak. Step 3: the non-linear least-squares routine is executed to meet the condition of NRMSE or DRMSE > ε. The original, calculated early-eluted, later-eluted peaks and the error curve are shown as solid (a), dotted (b) and (c) and dashed (d) lines, respectively.

The noise-added overlapped peak was then smoothed and differentiated numerically utilizing the Savitzky–Golay seven-point quadratic least-squares method [28,29]. The smoothed peak and the differential peak wera used as raw data for computation of starting points, maxima and end points of component peaks utilizing the Weimann algorithm [34]. The number of component peaks was then obtained by counting these sets. In order to extract later eluted component peak parameters, the range of the time pocket had to be established. This was done by choosing the maximum and its 20% height point in the tail part of an overlapped peak as initial starting and final time points, respectively. The initial starting point was then shifted 5 points stepwise toward the valley for each iteration until the minimum PRMSE was obtained. The later eluted peak parameters were then extracted in this optimized time pocket. The early-eluted peak parameters were obtained by subtracting these values from the overall peak and optimization process. The measure of the goodness of fit of the calculated peak to the original peak was tested utilizing the method of root mean square error of normal (NRMSE) and derivative peaks (DRMSE). If both calculations met an execution criterion, both NRMSE and DRMSE less than a value ε, our initial guessing routine was completed for peak deconvolution. The graphical representation of these procedures is illustrated in Fig. 1, steps 1–3, together with error curve, step 3. If the criteria were not met, the deconvolution was processed further utilizing a non-linear least-squares routine.

In all simulation procedures, statistical estimations such as mean values, confidence limit, reproducibility and relative percentage error of each parameter were calculated from twenty runs.

EXPERIMENTAL

Reagents

Nitrogen (99.95%) was used as the carrier gas after passing it through a molecular sieve. Ethylene (Matheson) and its deuterated isotope (Merck Sharp and Dohme of Canada) were used without further purification.

Column

A 30 m × 2 mm I.D. PTFE column packed with 228.6 mg of dicarbonylrhodium(I) 3-trifluoroacetyl-(1*R*)-camphorate (Johnson Matthey) in 9.0 g of squalane (Merck) coated on 52.9 g of Chromosorb P AW (30–60 mesh) was used to separate ethylene isotopes.

Apparatus

A Hewlett-Packard HP 5880A and a laboratory-made gas chromatograph equipped with flame ionization detectors were used. Samples were injected by use of a 1.0-ml internal loop gas sampling valve (Valco) and the pressure was monitored by use of a pressure transducer (Datametrics). The signal was taken as an averaged point every 6 s (equivalent to 180 points average) for the HP 5880A and every 3 s (equivalent to 30 points average) for the laboratory-made gas chromatograph.

RESULTS AND DISCUSSION

Simulation studies

Computer simulation studies were carried out to demonstrate the applicability of the present technique to real chromatograms. The resolution of asymmetric chromatograms is governed by the number of component peaks, noise, peak shape, peak size and degree of overlap. The number of component peaks in this study was limited to two for simplicity and to avoid any confounding effects, as mutli-component systems with more than three components are simply extensions of two-component systems (see also the end of this section). For the noise effect, a 1.0% noise level (corresponding to a maximum signal-to-noise ratio = 100), a possible worst case in a commercial chromatogram, was chosen for the studies as the noise level is indirectly proportional to the degree of recovery of peak parameters [9,26]. The peak shape was measured by observing the τ/σ value for a single-peak chromatogram. For an overlapped chromatogram, the peak shapes were studied in two distinct cases separately, *i.e.*, with constant τ and with variable τ. With constant τ, there could be two different cases of peak shapes. If the σ values of two components were the same, the peak shapes of both components were represented in terms of τ/σ and, if not ($\sigma_1 \neq \sigma_2$), by τ and σ_1/σ_2. With variable τ, the peak shape was represented mainly by τ_1/τ_2 as the effect of τ predominated over the σ values. The peak size effect was expressed by the peak-area ratio, A_1/A_2, of two components. For the measure of overlap we introduced a variance-independent reduced resolution (*RR*) as its description by conventional resolution, as pointed out by previous workers [35,36], is inadequate for an asymmetric peak. The *RR* was defined by a given resolution (a function of retention times and peak shapes) divided by the critical resolution (a function of peak shapes only), the value at the complete disappearance point of the valley between two component peaks. In order to obtain *RR*, the critical resolution had to be calculated. This could be done by bringing two component peaks (with τ/σ and A_1/A_2 known) from infinite separation to the disappearance point of the valley. The collection of all possible values then formed a contour map on the τ/σ and A_1/A_2 surface, as shown in Fig. 2. Simulated chromatograms are displayed in Fig. 3 as functions of τ/σ, A_1/A_2 and *RR*, obtained from the map.

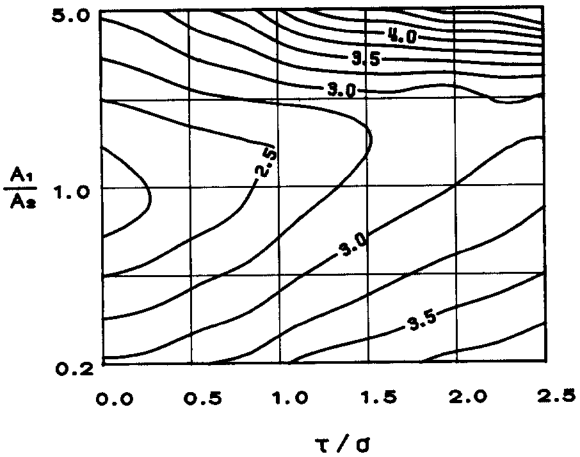

Fig. 2. Contour map of critical separation on the τ/σ and A_1/A_2 surface. Critical resolution = critical separation/total peak variance.

Case 1. Constant τ. The applicability of the technique was tested for various conditions and the results are listed in Tables I–III for constant ($\sigma_1 = \sigma_2$) and Table IV for different peak broadening parameters ($\sigma_1 \neq \sigma_2$). In Table I for a comparative study between initial guesses and non-linear regressions at $A_1 = A_2$ and Table II for intitial guessing routines as functions of *RR*, τ/σ and A_1/A_2, the relative percentage errors of extracted peak parameters from initial guesses showed decreasing tendencies from *RR* = 0.6, the complete merging point of two peaks, to *RR* = 1.4, from where conventional geometrical methods [1–4] could be applied with reasonable accuracies. In these regions the relative percentage errors were found to be <2% in most instances for *RR* from >1.0 to 1.4 applying an initial guessing routine alone, while the errors were increased drastically for 0.6 ⩽ *RR* ⩽ 1.0 where the accuracies could only be improved up to <2.2% by applying a non-linear least-squares routine. The study was stopped at *RR* = 0.6, the complete merging point of two peaks, as its lower limit, as it is impractical below this point (see Fig. 3). Through-

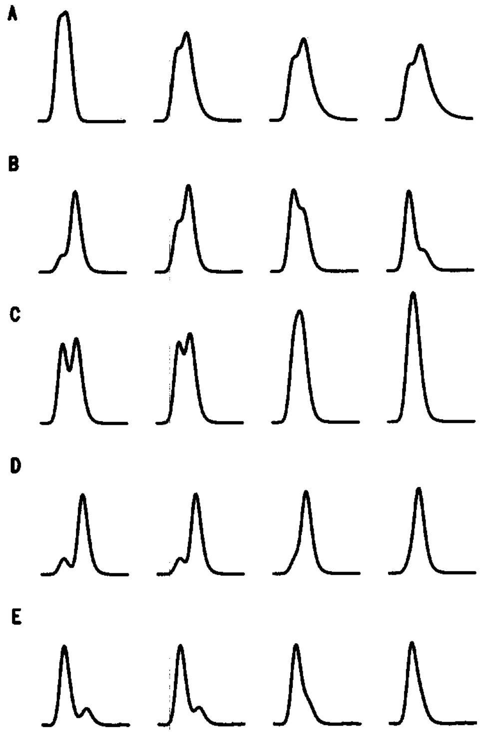

Fig. 3. Simulated chromatograms as a function of reduced resolution (RR), peak-area ratio (A_1/A_2) and asymmetry (τ/σ). (A) Chromatograms for $\tau/\sigma = 0.5$, 1.5, 2.0 and 2.5 at fixed $RR = 1.0$ and $A_1/A_2 = 1.0$. (B) Chromatograms for $A_1/A_2 = 0.2/1$, 0.6/1, 1/0.6 and 1/0.2 at fixed $RR = 1.0$ and $\tau/\sigma = 1.0$. (C) Chromatograms for $RR = 1.4$, 1.2, 0.8 and 0.6 at fixed $A_1/A_2 = 1.0$ and $\tau/\sigma = 1.0$. (D) Same conditions as (C), except $A_1/A_2 = 0.2/1$. (E) Same conditions as (C), except $A_1/A_2 = 1/0.2$.

out the studies it was also found that RR was a major factor for the accuracies of peak recovery processes because of its prominent effect compared with those of A_1/A_2 and τ/σ. The present technique showed a general tendency for underestimation for early-eluted peaks and overestimation for later-eluted peaks. These effects indicated a positive contribution from the tail part of an early-eluted peak to the later-eluted peak parameters and, in turn, a negative contribution to the early-eluted peak, obtained by subtracting the later-eluted peak from the convoluted peak. Hence the greater the later-eluted peak area was, the better was the recovery of peak parameters because it suffered relatively less perturbation from the early-eluted peak. The recovery was also more effective in the case of larger τ/σ values for given RR and A_1/A_2. This may be attributed to the fact that the overlapped peak with relatively larger τ/σ value has a larger retention time difference (Δt_G) and hence the obtained peak parameters from the resulting Gaussian component show better precision. In Table III, the relative percentage errors and reproducibility of the extracted peak parameters by non-linear regression at constant peak shape ($\sigma_1 = \sigma_2$) are found to be <2% in most instances and the worst values are 5.94 and 4.29%, respectively. In Table IV, for the parameters with different peak broadenings ($\sigma_1 \neq \sigma_2$), the errors show very minor differences from the case of constant τ/σ even for large variations of σ_1/σ_2 as long as τ remains constant. From the foregoing discussion, it may be concluded that the tailing parameter τ plays a major role in the extraction of peak parameters.

Case 2. Variable τ. Comparative studies were performed using a full EMG with variable τ and a simplified EMG with constant τ, and the results are listed in Table V. The relative percentage errors of t_G and A obtained by eqn. 3 were <2% for all calculations, whereas with eqn. 5 they increased by an average of 2.5% for every 10% difference in τ values ($\Delta\tau = |\tau_1 - \tau_2|$). The average computing times for these procedures were also compared and were 9 s for initial guessing 2 min for non-linear regression at constant τ and 20 min for non-linear regression at variable τ with an IBM-XT personal computer. Therefore, one may conclude that the simplified EMG calculation is more practical than the full EMG calculation for constant τ and still useful even for variable τ at the expense of reproducibility up to 5% when the difference in two τ values is less than 10%.

Although the detailed consideration of the technique was limited to a two-component system, the technique was also tested on a four-component system to see the possibility of extending it to multi-component systems. Deconvolution was performed using the last peak as a starting component and the remainder was thereby obtained by subtracting the calculated peak from the overall peak. The same procedure was repeated using the remainder until the whole peak was deconvoluted and the resolved peaks are displayed in Fig. 4. The relative percentage errors and reproducibilities were found to be <2.1 and <1.1%, respectively.

TABLE I

RELATIVE PERCENTAGE ERRORS OF EXTRACTED PEAK PARAMETERS WITH THE INITIAL GUESSING AND WITH THE NON-LINEAR LEAST-SQUARES ROUTINES AT VARIOUS *RR*

Data were obtained at fixed τ/σ = 1.0, A_1/A_2 = 1.0 and with added normal random noise of signal-to-noise ratio = 100 with respect to the maximum peak height.

RR	Peak[a]	Initial guessing				Non-linear least-squares			
		τ	σ	t_G	A	τ	σ	t_G	A
0.6	L	−5.52	1.75	−0.45	24.62	−2.10	2.15	0.07	0.92
	E	−62.48	−3.98	0.75	−32.45	—	−1.75	0.05	−1.04
0.8	L	−1.58	1.05	−0.16	9.25	−1.84	1.96	0.05	0.51
	E	−22.73	−2.43	0.20	−8.45	—	1.12	0.04	−0.84
1.0	L	−0.87	1.63	−0.08	3.27	−1.54	1.33	0.02	0.13
	E	−6.78	0.52	0.07	−3.49	—	1.15	0.01	−0.19
1.2	L	−0.52	0.65	−0.02	0.94				
	E	−2.45	0.46	0.05	−1.14				
1.4	L	−0.33	0.20	0.01	0.17				
	E	−1.05	0.10	0.14	−0.22				

[a] E and L denote the early- and the later-eluted peaks, respectively.

TABLE II

RELATIVE PERCENTAGE ERRORS OF EXTRACTED PEAK PARAMETERS WITH THE INITIAL GUESSING ROUTINE AT VARIOUS *RR*, τ/σ AND A_1/A_2

Data were taken from noise-added peaks. The noise was normal random noise of signal-to-noise ratio = 100 with respect to the maximum peak height.

RR	τ/σ	A_1/A_2	Later-eluted component				Early-eluted component			
			τ	σ	t_G	A	τ	σ	t_G	A
1.0	0.5	1/1	−7.35	2.63	0.27	6.57	−6.51	2.13	0.50	−4.25
	1.5	1/1	−0.36	1.95	−0.06	3.39	−1.33	−1.37	0.20	−3.96
	2.5	1/1	0.20	−0.56	−0.49	2.53	−0.06	−0.57	0.04	−2.33
1.2	0.5	1/1	−4.52	0.54	0.06	0.74	−3.09	0.36	−0.01	−0.96
	1.5	1/1	−0.31	0.45	−0.01	0.61	−1.82	0.35	0.03	−0.48
	2.5	1/1	0.62	−0.21	−0.01	0.45	0.14	−0.42	0.11	−0.18
1.0	1.0	0.2/1	−2.45	1.52	0.07	0.07	−8.45	1.62	0.04	−0.84
	1.0	0.6/1	−1.55	1.87	0.10	0.89	−3.23	2.03	0.15	−1.27
	1.0	1/0.6	−1.75	1.58	0.20	1.63	−2.75	1.38	0.15	−1.78
	1.0	1/0.2	−4.61	3.62	−0.33	5.09	−10.51	1.66	0.22	−2.23
1.2	1.0	0.2/1	−0.34	0.48	−0.00	0.05	−1.99	0.11	0.01	−0.23
	1.0	0.6/1	−0.13	0.11	−0.00	0.11	−1.31	0.13	0.02	−0.14
	1.0	1/0.6	−0.70	0.62	−0.01	1.02	−3.68	0.24	0.05	−1.08
	1.0	1/0.2	−1.35	1.46	−0.03	2.32	−6.53	0.91	0.08	−2.48

TABLE III

RELATIVE PERCENTAGE ERRORS AND REPRODUCIBILITIES OF EXTRACTED PEAK PARAMETERS WITH THE NON-LINEAR SQUARE ROUTINE AT VARIOUS RR, τ/σ AND A_1/A_2

Data were taken from noise-added peaks. The noise was normal random noise of signal-to-noise ratio = 100 with respect to the maximum peak height.

RR	τ/σ	A_1/A_2	Peak[a]	Relative error (%)				Reproducibility (%)			
				τ	σ	t_G	A	τ	σ	t_G	A
0.6	0.5	1/1	L	−5.94	1.82	0.21	1.09	4.21	0.92	0.24	1.42
			E	–	1.21	0.10	−0.95	–	1.32	0.13	1.59
0.6	1.5	1/1	L	−2.21	1.14	0.10	0.75	2.14	1.22	0.15	1.21
			E	–	0.48	0.14	−0.54	–	0.63	0.21	1.35
0.6	2.5	1/1	L	−3.29	1.54	−0.12	0.75	2.62	1.37	0.04	0.97
			E	–	−0.82	−0.17	−0.95	–	0.14	0.05	1.02
0.8	0.5	1/1	L	−2.07	0.52	0.08	0.63	1.49	0.31	0.00	0.16
			E	–	0.44	0.07	−0.72	–	0.46	0.00	0.10
0.8	1.5	1/1	L	−0.92	0.37	0.02	0.62	0.87	0.21	0.01	0.49
			E	–	0.51	0.02	−0.69	–	0.39	0.01	0.05
0.8	2.5	1/1	L	−1.44	0.21	−0.00	0.36	1.38	0.40	0.01	0.26
			E	–	0.21	−0.04	−0.55	–	0.15	0.01	0.31
0.6	1.0	0.2/1	L	−1.44	−2.63	1.82	−4.20	0.74	1.24	0.04	1.69
			E	–	1.16	0.42	2.33	–	0.10	0.25	1.51
0.6	1.0	0.6/1	L	−1.30	1.23	0.03	0.04	0.08	0.09	0.10	0.16
			E	–	1.14	0.02	−0.18	–	0.08	0.00	0.31
0.6	1.0	1/0.6	L	−1.54	1.38	0.04	0.05	0.23	0.25	0.01	0.37
			E	–	1.29	0.02	−0.18	–	0.80	0.00	0.19
0.6	1.0	1/0.2	L	−1.88	1.69	0.14	0.74	3.92	3.63	0.04	4.29
			E	–	2.18	0.04	−1.46	–	0.18	0.01	2.01
0.8	1.0	0.2/1	L	−0.17	−1.20	0.02	0.09	0.25	0.05	0.00	0.05
			E	–	1.27	0.14	−0.19	–	0.17	0.05	0.47
0.8	1.0	0.6/1	L	−1.27	1.18	0.04	−0.00	0.12	0.08	0.00	0.06
			E	–	1.15	0.04	−0.00	–	0.09	0.00	0.11
0.8	1.0	1/0.6	L	−1.52	1.31	0.04	0.11	0.31	0.22	0.00	0.23
			E	–	1.17	0.04	−0.10	–	0.08	0.00	0.15
0.8	1.0	1/0.2	L	−1.88	1.33	0.03	0.63	2.92	2.16	0.18	3.12
			E	–	1.19	0.42	−1.48	–	0.19	0.00	1.94

[a] E and L denote the early- and the later-eluted peaks, respectively.

Experimental verification

An experimental study was carried out to test the validity of the present algorithm using a PC-based data acquisition system. An isotopic mixture of [$^2H_2$]ethylene and [$^2H_3$]ethylene was chosen as a prototype of deconvolution problems among three typical cases such as isotopic mixture, enantiomer and isomer separations, utilizing the complexation gas chromatographic (GC) technique [37]. The reference values of peak parameters for individual components were obtained by injecting each sample into the GC system separately and utilizing our previous technique [26]. Premixed samples of known compositions of each component were then injected to test the algorithm. The GC conditions, such as carrier gas flow-rate and column temperature, were

TABLE IV

RELATIVE PERCENTAGE ERROR CHANGES OF t_G AND A AT CONSTANT τ/σ AND AT VARIOUS σ_1/σ_2

Data were obtained from noise-added peaks using the non-linear least-squares routine. The noise was normal random noise of signal-to-noise ratio = 100 with respect to the maximum peak height.

σ_1/σ_2	Later-eluted component		Early-eluted component	
	t_G	A	t_G	A
0.2/1	0.07	0.14	0.03	0.09
0.6/1	0.05	0.07	0.01	0.04
1/0.6	0.04	0.02	0.06	0.07
1/0.2	0.03	0.08	0.03	0.12

controlled to generate two different RR regions, *i.e.*, one with $RR > 1$ required the initial guessing routine only, and the other with $RR \leqslant 1$ required the non-linear regression. Averaged chromatograms from ten runs with the mixture samples under two different conditions and those of the two individual components are displayed in Fig. 5A for $RR > 1$ and Fig. 5B for $RR \leqslant 1$, as dotted lines together with the calculated values (open circles). The verified results of the present algorithm are presented in Table VI, with relative errors <1.52% under both conditions and standard deviations <2.14%. The present technique demonstrates reasonable agreement between the computed peak profiles from the overlapped peaks of mixtures and the measured peak profiles obtained from individual injections.

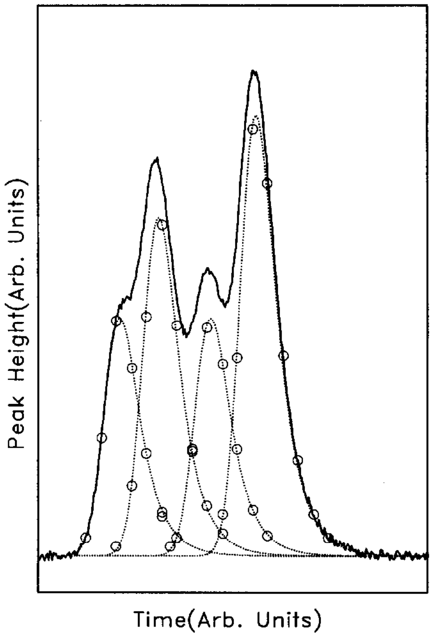

Fig. 4. Typical deconvolution from an overlapped four-component system. The overlapped chromatogram was synthesized from four sets of known peak parameters. The solid line, dotted lines and open circles are the overlapped peak, component peaks and deconvoluted component peaks utilizing this technique, respectively.

We also investigated the linearity of the peak area

TABLE V

RELATIVE PERCENTAGE ERRORS OF t_G AND A USING THE SIMPLIFIED EMG AND THE FULL EMG FUNCTIONS FOR VARIOUS τ_1/τ_2

Data were obtained from noise-added peaks using the non-linear least-squares routine. The noise was normal random noise of signal-to-noise ratio = 100 with respect to the maximum peak height.

τ_1/τ_2	Simplified EMG				Full EMG			
	Later-eluted component		Early-eluted component		Later-eluted component		Early-eluted component	
	t_G	A	t_G	A	t_G	A	t_G	A
0.8/1	−0.30	5.08	0.29	−5.13	0.05	0.84	0.04	−0.91
0.9/1	0.00	2.61	0.16	−2.65	0.07	0.72	0.02	−0.73
1/1	0.04	0.06	0.04	−0.10	0.03	0.12	0.05	−0.19
1/0.9	0.08	−2.59	−0.07	2.48	0.06	0.59	0.03	−0.72
1/0.8	0.12	−5.23	−0.16	5.10	0.04	1.21	0.07	−1.32

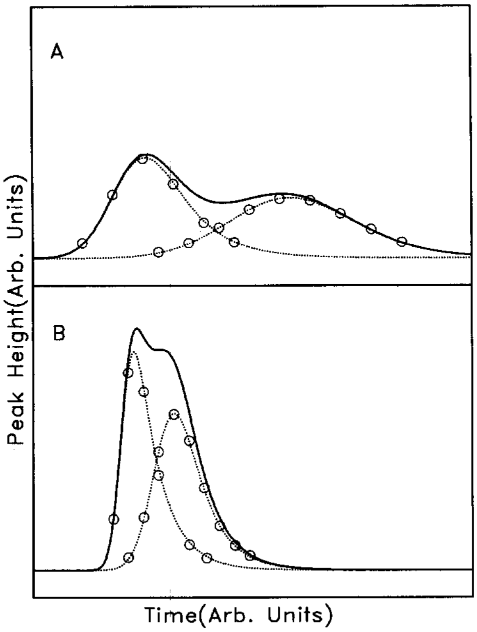

Fig. 5. Comparison between the observed and the deconvoluted chromatograms calculated from overlapped peaks for two distinct cases, *i.e.*, (A) for $RR > 1$ and (B) for $RR \leqslant 1$. (A) For $RR > 1$: the dotted lines are obtained by injecting two individual samples separately and the solid line by injecting the mixture of two components with GC conditions of carrier gas flow-rate 17 ml/min, oven temperature 20°C and samples [$^2H_2$]ethylene and [$^2H_3$]ethylene. The open circles represent the resolved peak points from the chromatogram of the mixture by the present algorithm. (B) For $RR \leqslant 1$: the lines and the open circles are same in (A) except for the GC conditions: carrier gas flow-rate 25 ml/min, oven temperature 60°C and samples [$^2H_2$]ethylene and [$^2H_3$]ethylene.

TABLE VII

RECOVERED PEAK AREAS USING THE PROPOSED TECHNIQUE

Sample used (Torr)		Recovered peak area[a]	
[$^2H_2$]	[$^2H_3$]	[$^2H_2$]	[$^2H_3$]
0.201	0.401	1.00	2.03
0.398	0.197	1.94	1.01
0.604	0.605	2.96	3.09

[a] The chromatographic peak areas recovered by the proposed method are represented as reduced peak areas with respect to [$^2H_2$]ethylene.

to test the applicability of this method in quantitative analysis, and the results are presented in Table VII. The observed area of the later-eluted peak of an equimolar mixture of two components was always greater than that of the early-eluted peak, and these results were also consistent with those of our simulation studies. The correction factors of the computed values against the initially known amounts of samples were found to be 0.982 for [$^2H_2$]ethylene and 1.006 for [$^2H_3$]ethylene. These findings also indicate that the present technique is suitable for the quantitative analysis of overlapped components.

TABLE VI

RELATIVE PERCENTAGE ERRORS AND STANDARD DEVIATIONS OF EXTRACTED PEAK PARAMETERS FROM EXPERIMENTAL CHROMATOGRAMS

Conditions[a]	Peak[b]	Relative error (%)[c]				Standard deviation (%)			
		τ	σ	t_G	A	τ	σ	t_G	A
A	L	−0.27	0.47	−0.01	0.38	1.32	1.07	0.03	0.51
	E	−1.13	0.20	0.03	−0.43	2.14	1.45	0.07	0.74
B	L	−0.10	0.13	−0.00	0.01	1.84	1.42	0.07	0.98
	E	−0.07	−0.04	0.00	−0.00	—	1.64	0.25	1.24

[a] Conditions A: the acquired conditions are $RR = 1.27$, $\tau_1/\tau_2 = 0.97$, $\sigma_1/\sigma_2 = 0.36$ and $A_1/A_2 = 1.00$. Peak parameterization requires the initial guessing routine only (see also Fig. 5A). Conditions B: the acquired conditions are $RR = 0.94$, $\tau_1/\tau_2 = 0.96$, $\sigma_1/\sigma_2 = 0.40$ and $A_1/A_2 = 1.00$. Peak parameterization requires the non-linear least-squares method using simplified EMG (see also Fig. 5B).

[b] E and L denote the early-eluted ([$^2H_2$]ethylene) and the later-eluted component ([$^2H_3$]ethylene), respectively.

[c] Relative percentage error of peak parameters, P, is defined by

$$P = \frac{(P_{\text{sep}} - P_{\text{mix}})}{P_{\text{sep}}} \cdot 100$$

where P_{sep} and P_{mix} are the peak parameters of a single peak obtained by the method of ref. 26 and those of the convoluted peak, respectively.

ACKNOWLEDGEMENTS

This work was done with financial assistance from the Korea Standards Research Institute and the Korea Science and Engineering Foundation, which is gratefully acknowledged.

REFERENCES

1 J. D. Hettinger, J. R. Hubbard, J. M. Gill and L. A. Miller, *J. Chromatogr. Sci.*, 9 (1971) 710.
2 H. A. Hancock, Jr., L. A. Dahm and J. F. Muldoon, *J. Chromatogr. Sci.*, 8 (1970) 57.
3 A. W. Westerberg, *Anal. Chem.*, 41 (1969) 1770.
4 A. N. Papas and T. P. Tougas, *Anal. Chem.*, 62 (1990) 234.
5 R. F. Lacey, *Anal. Chem.*, 58 (1986) 1404.
6 A. H. Andreson, T. C. Gibb and A. B. Littlewood, *J. Chromatogr. Sci.*, 8 (1970) 640.
7 R. A. Vaidya and R. D. Hester, *J. Chromatogr.*, 287 (1984) 231.
8 M. Rosenbaum, V. Hancil and R. Komers, *J. Chromatogr.*, 191 (1980) 157.
9 K.-H. Jung and Y. H. Shin, *Bull. Korean Chem. Soc.*, 7 (1986) 403.
10 C. F. Lam, A. Forst and H. Bank, *Appl. Spectrosc.*, 33 (1979) 273.
11 R. A. Caruana, R. B. Searle, T. Heller and S. I. Shupack, *Anal. Chem.*, 58 (1986) 1162; R. A. Caruana, R. B. Searle and S. I. Shupack, *Anal. Chem.*, 60 (1988) 1896.
12 A. S. Said, *Theory and Mathematics of Chromatography*, Hüthig, Heidelberg, Basle, New York, 1981, p. 69.
13 K. Yamaoka and T. Nakagawa, *Anal. Chem.*, 47 (1975) 2050.
14 A. Jaulmes and C. Vidal-Madjar, *Anal. Chem.*, 63 (1991) 1165.
15 V. Maynard and E. Grushka, *Anal. Chem.*, 44 (1972) 1427.
16 A. B. Littlewood, *Gas Chromatography*, Academic Press, New York, 2nd ed., 1970, p. 169.
17 I. G. McWilliam and H. C. Bolton, *Anal. Chem.*, 41 (1969) 1755.
18 I. G. McWilliam and H. C. Bolton, *Anal. Chem.*, 41 (1969) 1762.
19 I. G. McWilliam and H. C. Bolton, *Anal. Chem.*, 43 (1971) 883.
20 T. S. Buys and K. de Clark, *Anal. Chem.*, 44 (1972) 1273.
21 E. Grushka, M. N. Myers and J. C. Giddings, *Anal. Chem.*, 42 (1970) 21.
22 O. Grubner, *Anal. Chem.*, 43 (1971) 1934.
23 R. E. Pauls and L. B. Rogers, *Anal. Chem.*, 49 (1977) 625.
24 J. P. Foley, *Anal. Chem.*, 59 (1987) 1984.
25 J. P. Foley and J. G. Dorsey, *J. Chromatogr. Sci.*, 22 (1984) 40.
26 K.-H. Jung, S. J. Yun and S. H. Kang, *Anal. Chem.*, 56 (1984) 457.
27 R. Kolvoda, *Operation Amplifiers in Chemical Instrumentation*, Wiley, New York, 1975, p. 49.
28 P. A. Gorry, *Anal. Chem.*, 62 (1990) 570.
29 A. Savitzky and M. J. E. Golay, *Anal. Chem.*, 36 (1964) 1672.
30 D. W. Maquardt, *J. Soc. Ind. Appl. Math.*, 11 (1963) 431.
31 W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, 1986, p. 523.
32 S. J. Yun, *Ph. D. Thesis*, Korea Advanced Institute of Science and Technology, Seoul, 1987.
33 W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, 1986, p. 202.
34 B. Weimann, *Chromatographia*, 7 (1974) 472.
35 J. J. Kirkland, W. W. Yau, H. J. Stoklosa and C. H. Diks, Jr., *J. Chromatogr. Sci.*, 15 (1977) 303.
36 J. P. Foley, *J. Chromatogr.*, 384 (1987) 301.
37 V. Schurig, *Chromatographia*, 13 (1980) 263.