

CHROM. 16,344

APPLICABILITY LIMITS OF THE EDGEWORTH-CRAMÉR SERIES IN CHROMATOGRAPHIC PEAK SHAPE ANALYSIS

FRANCESCO DONDI* and FERNANDO PULIDORI

Analytical Chemistry Laboratory Department of Chemistry Via L. Borsari 46, I-44100 Ferrara (Italy)

(Received October 5th, 1983)

SUMMARY

A simulation study on the properties of chromatographic peak shape data-handling software based on the Edgeworth-Cramér series is described. The least-squares approximation properties and the peak parameter bias and precision are analysed as a function of peak shape, noise and baseline determination. Applicability limits of the method and the maximum useful expansion order are defined. Up to a maximum peak skewness (S) value of 0.7, the mean, the standard deviation, the skewness and the excess can be determined with a precision interval lower than the bias, under normal noise conditions. In the range $0.7 < S < 1.0$, the parameter bias is slightly greater than the precision. S values as low as 0.05 can be determined. Optimum values of the signal-to-noise ratio are required for measuring the peak excess. Peak shapes with $S > 1.0$ cannot be handled in this way.

INTRODUCTION

In previous papers¹⁻³ concerning chromatographic peak shape analysis, the validity and the practical usefulness of the Edgeworth-Cramér (EC) series were examined. In order to establish the usefulness and define the practical employment of these series expansion, three questions must be answered. The first is the greatest peak skewness that can be handled. The second is the optimum expansion order to be used in fitting a given peak under determined conditions of detector noise. The third concerns the peak parameter precision and accuracy and their dependence on series truncation, signal-to-noise ratio and baseline determination. In fact, the peak parameters that can be determined by these series (mean, standard deviation, skewness, excess, higher cumulant coefficients)¹ are meaningful physical quantities, related to the mass transfer process inside the column and to extra-column band broadening phenomena⁴⁻⁸.

The aim of this work was to clarify the above points with respect to nearly symmetrical or moderately skewed peaks, which are often found in chromatographic practice.

The method employed is the classical simulation procedure⁹⁻¹¹. The principal interest is in the accuracy, and much importance is attached here to the difference

between the true and computed quantities. A great advantage of the curve fitting method is that an estimate of parameter precision is easily obtained. Thus attention is devoted to establishing the conditions under which the precision interval covers the bias. The essential advantage lies in the fact that, when used under the above tested conditions, the method can be considered as an unbiased method of peak parameter determination.

THEORETICAL

A chromatographic peak $Y(t)$ may be considered a standardized frequency function $f(x)$ with normalized unit area:

$$f(x) = Y(t) \sigma / A \quad (1)$$

where t is the time, σ the peak standard deviation, m the peak mean, A the area and x the standardized variable:

$$A = \int_{-\infty}^{+\infty} Y(t) dt \quad (2)$$

$$x = (t - m) / \sigma \quad (3)$$

The EC series expansion related to $f(x)$ and developed up to the k th order is

$$y(x) \approx Z(x) + \sum_{i=1}^{i=k} Q_i(-Z) \quad (4)$$

where $Z(x)$ is the normal frequency function:

$$Z(x) = 1/\sqrt{2\pi} \exp(-x^2/2) \quad (5)$$

and $Q_i(-Z)$ are linear aggregates in the derivatives of the normal frequency function $Z(x)$, of maximum order $3i$, containing cumulant coefficients of $f(x)$. The first two terms are

$$Q_1(-Z) = \frac{S Z^{(3)}(x)}{6} \quad (6)$$

and

$$Q_2(-Z) = \frac{E}{24} \cdot Z^{(4)}(x) + \frac{S^2}{72} \cdot Z^{(6)}(x) \quad (7)$$

where $Z^{(i)}(x)$ is the i th derivative of the normal frequency function $Z(x)$ with respect to x , and S and E are the peak skewness and excess, respectively^{1,2}.

The general expression, which allows us to compute to the tenth series term, is derived in Appendix B in ref. 1. In this paper a theoretical chromatographic peak shape arising from the stochastic theory^{12,13} of chromatography is considered:

$$f(t) = \exp(-\lambda_2 t - \mu) (\lambda_2 \mu / t)^{1/2} I_1(4\lambda_2 \mu t)^{1/2} \quad (8)$$

where μ is the mean number of adsorption-desorption steps, λ_2 is the time constant on the adsorption site and $I_1(x)$ is the modified Bessel function of the first kind.

COMPUTATIONAL PROCEDURE

A standard pseudo random Gaussian noise sequence, generated by computer, was added to the peak (eqn. 8) and no attempt was made to analyse the effect of the noise frequency spectrum or other error sources. A standard range of $\pm 10\sigma$ with respect to the peak mean and only positive time values were considered. The time increment was one tenth of the peak standard deviation. Identical results for both the minimum position and the peak parameter values were obtained by using different random noise sequences or denser sampling (up to $\sigma/100$), as observed in handling experimental peaks¹. The same general non-linear least-squares minimization procedure, as previously described (MINUIT)¹, was employed in fitting these peak profiles by EC series expanded up to the seventh order. The unknown peak parameters of the series are area A , peak mean m , variance σ^2 , peak skewness S , peak excess E and higher cumulant coefficients according to the series expansion order k ^{1,2}.

The peak parameter precision was computed on the basis of a 3 s_k increase in the minimum total standard deviation¹, where s_k is the root mean square error. Baseline determination was performed within the same minimization procedure, by considering an unknown parameter in the closed interval of ± 5 times the standard deviation of the random noise added to the peak, centred around the true baseline. More extended intervals gave the same results, but with longer computation times. All the calculations were run on a CDC Cyber 76 computer (CINECA, Casalecchio, Bologna, Italy).

The signal-to-noise ratio (S/N) is the ratio of the peak maximum to four times the standard deviation of the noise sequence. The standard deviations and the areas of all the theoretical peaks were the same ($\sigma = 10$ and $A = 10,000$). The degree of the fitting (approximation error, %) is expressed as the coefficient of variation [C.V. (%)] referred to the peak maximum, Y_{\max} :

$$\text{C.V.(\%)} = 100s_k / Y_{\max}$$

RESULTS AND DISCUSSION

Choice of the model

The peak shape function employed was derived from the stochastic theory of chromatography according to Giddings and Eyring¹² and McQuarrie¹³.

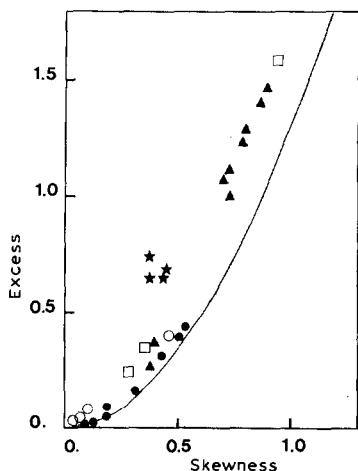


Fig. 1. Skewness vs. excess. The continuous line corresponds to (S/E) values analysed in this work. The points refer to experimental peaks taken from the literature: \square , ref. 1; \bullet , ref. 6; \circ , ref. 4; \star , peak profiles kindly supplied by Vidal-Majar and Guiochon¹⁵; \blacktriangle , ref. 16.

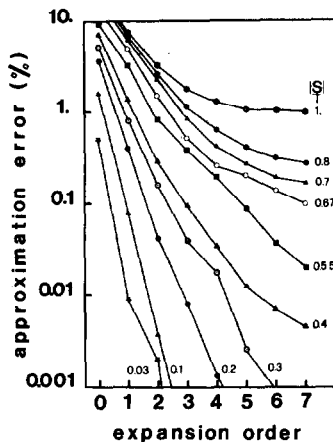


Fig. 2. Relative approximation error (C.V., %) as a function of the EC series expansion order and the peak skewness. Non-noisy profiles.

By a convenient choice of model parameters², peak shapes having the same standard deviation ($\sigma = 10$), but skewness values ranging from 0.03 to 1.0, have been generated. The excess skewness values that can be analysed in this way are reported in Fig. 1. In the same graph, (S,E) points referring to experimental gas chromatographic peaks, taken from various sources and determined by EC series^{1,14-16} or Gram-Charlier series⁶, are reported for sake of comparison. The low skewness values refer to packed and efficient gas-liquid chromatographic columns, and those of high skewness to physico-chemical studies by gas-solid chromatography. The type of theoretical peak shapes that were analysed in this work do not seem to differ greatly from those in chromatographic practice.

The bias and reproducibility of the peak shape analysis are functions of random error and of systematic error produced by the instrumentation and the data handling software. Interaction effects are possible between random noise and the data handling software⁴, so non-noisy profiles are considered separately.

Study on non-noisy profiles

The use of a higher expansion order generally improves the peak approximation, as expected. As observed in handling experimental peaks¹, the approximation degree is equally good on the entire peak and the deviations show an increasing number of nodes (change from + to -) by increasing the expansion order of the series. In Fig. 2, where the coefficient of variation is reported as a function of the expansion order of the EC series, very symmetrical peaks (up to a maximum S value of ca. 0.4) are approximated to better than 0.01%. Moderately skewed peaks are approximated to a lower precision. Very skewed peaks ($S = 0.8-1.0$) are approximated no more precisely than 0.3-1%. This last fitting level is inferior to a good experimental precision^{1,6,17}. In addition, for the most skewed peak ($S = 1.0$), no

TABLE I
PEAK PARAMETER RECOVERY ERROR (%) ON NON-NOISY PROFILES

Mean results obtained with EC series expansion orders greater than 2.

Parameter	$0.03 < S < 0.6$	$0.6 < S < 0.75$	$0.75 < S < 1.0$
A	<0.05	0.3-0.4	0.5
m^*	<0.06	0.3-1.7	0.3-2.5
σ^2	<0.5	0.4-1.5	1.5-3.0
S	<0.5	0.5-10	10-15
E	<15	10-20	20-30
Higher cumulant coefficients	>100	>100	>100

* The error (%) is expressed relative to the peak standard deviation.

improvement is achieved on passing from an expansion order of 5 to 7. This seems to be, in effect, the maximum limit of the fitting properties of the EC series, connected to the onset of the so-called asymptotic properties^{2,4,18-20}.

Surprisingly, we observed that the peak parameter recovery does not always improve with the expansion order, but sometimes greater accuracy is observed at lower expansion orders (e.g., $k = 2$ for σ^2 ; $k = 1$ for S ; $k = 2.4$ for E). The explanation of this findings is difficult, but probably lies in the correlation within the family of parameters of the same parity (odd: m, S, \dots ; even: σ^2, E, \dots)^{1,4,14}. In Table I, only the recovery error ranges as a function of the peak shape are reported. Very low systematic errors are observed for the main peak parameters A, m and σ^2 . Very low systematic errors are also found in measuring S of nearly Gaussian peaks. The peak excess E can be measured but less significantly. Higher cumulant coefficients cannot be determined at all.

Studies on noisy profiles

The superimposition of a random noise on the peak profile limits the maximum fitting precision attainable on expanding the series (Fig. 3). By using the F -test cri-

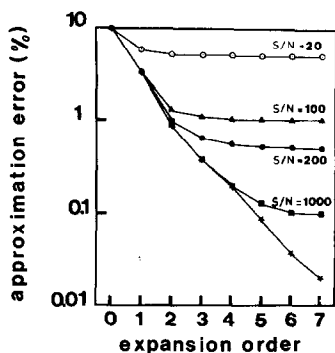


Fig. 3. Relative approximation error (C.V., %) as a function of the EC series expansion order and the signal-to-noise ratio. ★, Non-noisy profile. Peak skewness: 0.3.

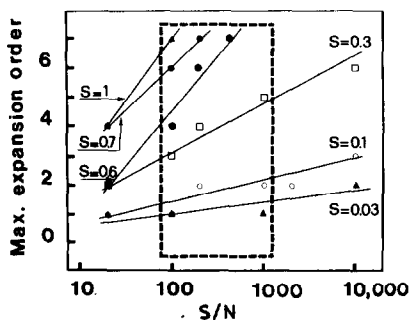


Fig. 4. Maximum expansion order of the best fitting EC series as a function of the signal-to-noise ratio and the peak shape.

TABLE II
PEAK PARAMETER RECOVERY ERROR (%) ON NOISY PROFILES

Parameter	$S/N = 100$		$S/N \cdot 1000$	
	$0.03 < S < 0.7$	$0.7 < S < 1.0$	$0.03 < S < 0.7$	$0.7 < S < 1.0$
A	0.2	0.2-0.8	0.05	0.2-0.7
m^*	0.3	0.4-1.0	0.2	0.3-0.7
σ^2	0.3	0.5-3.0	0.1	0.5-3.0
S	5-1	2-5	1-1.5	1-3
E	40-5.0	10-5.0	10-2.0	5
Baseline**	0.05	0.05-0.1	0.0025	0.025

* The error (%) is expressed relative to the peak standard deviation.

** The relative error is referred to the peak maximum.

terion, the maximum expansion order, k_{\max} , as a function of S/N can be determined for each peak shape¹. Thus Fig. 4 was built up, in which each line refers to a given peak skewness and the dotted area delimits the usual S/N values found when working with a flame-ionization detector¹⁷. Only for intermediate peak skewness values is a strong dependence between k_{\max} and S/N observed. Working conditions corresponding to the left-hand side of the dotted area signify bad experimental conditions, with a great loss of information on the peak properties. An increase in S/N beyond 1000 does not always mean a significant improvement in k_{\max} .

By extrapolating the pattern shown in Fig. 4, some inferences can be made as to the need for higher expansion orders beyond seven, which is the maximum expansion order employed here.

An increase in k_{\max} beyond 7 is expected to be significant only for high skewness values (0.8-1.0), when S/N has an optimum value (ca. 500)⁶. However, for this peak skewness the asymptotic limits of the EC series begin to appear, as discussed above. By increasing the expansion order by one unit, the natural consequence is that the attainable improvements are progressively less significant. Thus it can be concluded that the seventh EC expansion order seems to be sufficient for chromatographic applications, at least for low peak skewnesses.

The results in Figs. 3 and 4 refer to noisy profiles with the exact baseline. However, nearly identical results with only minor changes ($k_{\max} \pm 1$ in Fig. 4) are obtained by considering the baseline as an additional unknown parameter. Below, where the peak parameter determination is discussed, only results obtained by considering the baseline as an unknown are reported. In addition, among the possible sets of peak parameters, we consider only the set resulting from the best fitting, that is, that which is obtained with an expansion order equal to k_{\max} . In Table II the parameter recovery errors are reported as relative percentages. At low peak skewness values ($0.03 < S < 0.7$), the relative accuracy is good and nearly constant. For $0.7 < S < 1.0$, a general increase in the systematic error is observed. Improvements in S/N almost always results in better accuracy. For the last parameter, E , a greater improvement in accuracy is observed on increasing S/N . Thus, if a detailed description of the peak shape is required and E is also to be determined, S/N must be adjusted to its highest level.

TABLE III

PEAK PARAMETER PRECISION CALCULATED ON THE $3s$, INCREASE OF THE MINIMUM OF THE SQUARED DEVIATIONS (NOISY PROFILES) (MEAN QUANTITIES)

Parameter	$S/N = 100$	$S/N = 1000$
A	$\pm 0.2\%$	$\pm 0.05\%$
m^*	$\pm 0.4\%$	$\pm 0.1\%$
σ^2	$\pm 1.5\%$	$\pm 0.25\%$
S^{**}	± 0.01	± 0.0025
E^{**}	± 0.015	± 0.005
Baseline ^{***}	$\pm 0.05\%$	$\pm 0.025\%$

* The error (%) is expressed relative to the peak standard deviation. The common peak standard deviation is 10.

** The errors on these quantities are expressed as absolute errors.

*** The relative error is referred to the peak maximum.

Let us consider now the peak parameter precision (Table III). As a general rule, the parameter precision intervals were found to be nearly independent of the actual parameter values and thus of the peak shape, the major dependence being once again on S/N .

By comparing the peak precision with the accuracy (Tables II and III), it appears that when the peak is moderately skewed ($0.03 < S < 0.7$), the dominant effect is the precision. In this instance the precision interval covers the true value of the quantity. Instead, when the peak skewness is high ($0.7 < S < 1.0$), the dominant effect is the accuracy, and thus the precision interval can result in an underestimate of the bias (make use of Fig. 1 to calculate the relative error on S and E).

TABLE IV

COMPARISON BETWEEN PARAMETER PRECISION DETERMINED ON THEORETICAL AND EXPERIMENTAL PEAKS BY EC SERIES

Signal-to-noise ratio = 100. Results are relative errors (%).

	<i>Low-skewness peaks</i>		<i>High-skewness peaks</i>	
	<i>Theoretical peak shape</i> ($S = 0.3$; $E = 0.12$)	<i>Experimental peak shape</i> ($S = 0.27$; $E = 0.32$)	<i>Theoretical peak shape</i> ($S = 0.89$; $E = 1.03$)	<i>Experimental peak shape</i> ($S = 0.95$; $E = 1.58$)
K_{max}	4	3	5	6
A	0.20	0.16	0.38	1.2
m^*	0.4	0.3	0.4	1.0
σ^2	0.53	0.77	1.5	3.1
S	4.9	7.7	2.7	4.2
E	24	11	6.6	3.8
Baseline ^{**}	0.067	0.025	0.09	0.03
Approx. error (%)	0.25	0.14	0.31	0.50

* The error (%) is expressed relative to the peak standard deviation.

** The relative error is referred to the peak maximum.

Comparison with experimental peaks

Two peaks, recorded as described in ref. 1, were used. They were derived from cases I and III discussed in ref. 1, but the baseline was treated as an additional unknown. No repeated peaks (by using the same sample) were considered, as it has been shown¹ that repeatability is the same as the precision computed on the $3 s_k$ basis using a single peak.

The results are given in Table IV. Excellent agreement is observed between the precision measured on experimental and theoretical peaks, the k_{\max} employed and the degree of fitting attained, at comparable levels of peak skewness.

CONCLUSIONS

The present data handling software appears to be superior to other peak treatments based on different peak shape functions^{1,6,21,22} or on the integration method^{9,11,23}, provided that the peak skewness is not greater than 1.0. In fact, the present method is valuable both in peak fittings and in peak parameter determinations. Previous methods fail in one or other of these aims⁶. For skewness values higher than 1.0, the limits inherent in the asymptotic properties of this series call for new peak shape functions^{4,20} or different data handling methods.

The conclusions on the validity of the EC series can, however, be applied chiefly to those asymmetry effects which are observed under conditions of infinite dilution and classified as site heterogeneity, mixed retention, poorly packed columns or slow retention kinetics²⁴⁻²⁶. In fact, both the peak shape function and the approximation function employed in this work come from the stochastic theory of the sums of random and independent variables. In chromatographic terms, they describe the trajectories of the solute molecule through the column only when they behave independently of each other^{12,13,19}. Care must be taken, instead, in dealing with the concentration-dependent asymmetry effects, which are known to be another important source of peak asymmetry in chromatography²⁷.

ACKNOWLEDGEMENT

This work was supported by the Italian Education Department (MPI) and was presented in part at the Fourth National Congress on Analytical Chemistry, Urbino, Italy, September 6-9th, 1982.

REFERENCES

- 1 F. Dondi, A. Betti, G. Blo and C. Bigli, *Anal. Chem.*, 53 (1981) 496.
- 2 F. Dondi, *Anal. Chem.*, 54 (1982) 473.
- 3 A. Betti, F. Dondi, G. Blo, G. Coppi, G. Cocco and C. Bigli, *J. Chromatogr.*, 259 (1983) 433.
- 4 P. C. Kelly and W. E. Harris, *Anal. Chem.*, 43 (1971) 1170.
- 5 O. Grubner, *Advan. Chromatogr.*, 6 (1968) 173.
- 6 C. Vidal-Madjar and G. Guiochon, *J. Chromatogr.*, 142 (1977) 61.
- 7 J. C. Sternberg, *Advan. Chromatogr.*, 2 (1966) 205.
- 8 E. Grushka, in N. Catsimpoolas (Editor), *Methods of Protein Separation*, Vol. 1, Plenum Press, New York, 1975, Ch. 6.
- 9 M. Goedert and G. Guiochon, *J. Chromatogr. Sci.*, 11 (1973) 326.
- 10 M. Goedert and G. Guiochon, *Chromatographia*, 6 (1973) 116.

- 11 T. Pettitclerc and G. Guiochon, *Chromatographia*, 7 (1974) 10.
- 12 J. C. Giddings and H. Eyring, *J. Phys. Chem.*, 59 (1955) 416.
- 13 D. A. McQuarrie, *J. Chem. Phys.*, 38 (1963) 437.
- 14 P. C. Kelly and W. E. Harris, *Anal. Chem.*, 44 (1972) 502.
- 15 C. Vidal-Madjar and G. Guiochon., unpublished results.
- 16 A. Betti, F. Dondi, G. Blo and C. Bigli, unpublished results.
- 17 D. H. Smith, *Advan. Chromatogr.*, 12 (1975) 177.
- 18 V. A. Kaminskii, S. F. Timashev and N. N. Tunitskii, *Russ. J. Phys. Chem.*, 39 (1965) 1354.
- 19 H. Cramér, *Random Variables and Probability Distributions*, Cambridge University Press, Cambridge, 1962, p. 70.
- 20 D. L. Wallace, *Ann. Math. Stat.*, 29 (1958) 635.
- 21 S. M. Roberts, D. H. Wilkinson and R. L. Walker, *Anal. Chem.*, 42 (1970) 886.
- 22 E. Grushka and S. D. Mott, *J. Chromatogr.*, 186 (1979) 117.
- 23 S. N. Chesler and S. P. Cram., *Anal. Chem.*, 43 (1971) 1922.
- 24 L. R. Snyder and J. J. Kirkland, *Introduction to Modern Liquid Chromatography*, Wiley-Interscience, New York, 1979, p. 791.
- 25 J. C. Giddings, *Anal. Chem.*, 35 (1963) 1999.
- 26 F. Dondi, in M. Kivimäki and L. Nünistö (Editors), *Euroanalysis IV Abstracts, Helsinki/Espoo, August 23-28, 1981*. Association of Finnish Chemical Societies, Helsinki, 1981, p. 374.
- 27 J. R. Conder and C. L. Young, *Physicochemical Measurement by Gas Chromatography*, Wiley, Chichester, 1979, p. 25.