

Improved computer algorithm for characterizing skewed chromatographic band broadening

I. Method

W. W. YAU* and J. J. KIRKLAND

E. I. du Pont de Nemours and Company, Central Research and Development Experimental Station, B-228, P.O. Box 80228, Wilmington, DE 19880-0228 (USA)

ABSTRACT

A method of extracting column band-broadening parameters from skewed and noisy chromatographic peaks is derived from an exponentially modified Gaussian peak model. In the proposed method of determining peak variance and peak skew, only four peak parameters need to be measured from an experimental chromatographic peak: peak retention time, peak height, peak area and the first moment or center of gravity of the peak. This proposed method is more accurate and less susceptible to baseline noise than those previously described.

INTRODUCTION

Extracting accurate information from real chromatographic peaks is a challenging task that requires sophisticated computational approaches. It is well known that use of the Gaussian peak model gives significant errors in determining plate number, peak asymmetry factor and resolution [1]. Because of this, various studies have proposed many approaches, the most accepted being based on the exponentially modified Gaussian (EMG) model [2–6]. Uses of the EMG model have been reviewed [7]^a, and the effect of random noise on the accuracy of measurements by the EMG model has been presented [8].

This paper presents a new method of extracting information from skewed and noisy peaks, based on the EMG model. This method is more accurate than previous approaches; susceptibility to errors from baseline noise is also substantially reduced. The method has been used with excellent results for more than 4 years in our laboratory for many high-performance liquid chromatographic, size-exclusion chromatographic and field-flow fractionation studies.

^a A review of EMG function since 1983 has just been published [13].

THEORY

The following EMG peak-shape model is used:

$$h(t) = \left(\frac{A}{2\tau}\right) \exp\left(\frac{\sigma^2}{2\tau^2} - \frac{t - t_R}{\tau}\right) [1 + \operatorname{erf}(Z/\sqrt{2})] \tag{1}$$

where

$$Z = \frac{t - t_R}{\sigma} - \frac{\sigma}{\tau} \tag{2}$$

and

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy \tag{3}$$

Fig. 1 depicts the EMG peak-shape model as described by eqn. 1, where h_p is the peak height, t_p is the peak retention time, t_R is the retention time of the Gaussian input function, σ is the standard deviation of the Gaussian component, τ is the time constant of the exponential modifier and M_1 is the first moment of the skewed experimental peak. The M_1 value denotes the centroid retention time of the skewed experimental peak. Peak moments are defined by

$$\text{peak area } (A) = M_0 = \int_{\text{all } t} h(t) dt \tag{4}$$

$$\text{peak centroid} = M_1 = \frac{\int_{\text{all } t} th(t) dt}{M_0} \tag{5}$$

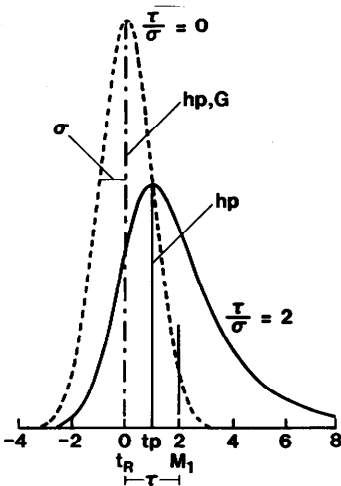


Fig. 1. Peak shape model: exponentially modified Gaussian.

With this peak-shape model, the following useful properties of peaks are described [3,9,10]:

$$M_1 = t_R + \tau \quad (6)$$

$$\text{peak variance} = \sigma^2 + \tau^2 \quad (7)$$

$$\text{peak skew} = Sk = 2\tau^3/(\sigma^2 + \tau^2)^{3/2} \quad (8)$$

$$\text{peak excess} = (3\sigma^4 + 6\sigma^2\tau^2 + 9\tau^4)/(\sigma^2 + \tau^2)^2 - 3 \quad (9)$$

and

$$\text{plate number } (N) = (t_R + \tau)^2/(\tau^2 + \sigma^2) \quad (10)$$

From the above equations, other useful properties of peaks can be defined:

$$h_p = \frac{A}{\sigma\sqrt{2\pi}} \cdot \exp(-\gamma^2/2) \quad (11)$$

where

$$\gamma = \text{peak displacement} = (t_p - t_R)/\sigma \quad (12)$$

and

$$\frac{\tau}{\sigma} = \sqrt{\pi/2} \exp(Z_p^2/2)[1 + \text{erf}(Z_p/\sqrt{2})] \quad (13)$$

where

$$Z_p = Z\text{-value at peak} = (t_p - t_R)/\sigma - \sigma/\tau \quad (14)$$

The τ/σ and the γ values for a broadened peak of a particular Z_p value can be calculated from eqns. 12–14, as exemplified in Table I.

To conserve computer time, the $\text{erf}(Z_p/\sqrt{2})$ in eqn. 13 can be calculated with the following approximation:

$$\text{erf}(Z_p/\sqrt{2}) = \text{erf}(x), \quad \text{for } Z_p > 0 \quad (15a)$$

$$\text{erf}(Z_p/\sqrt{2}) = -\text{erf}(x), \quad \text{for } Z_p < 0 \quad (15b)$$

where

$$x = |Z_p/\sqrt{2}| = \text{absolute value of } Z_p/\sqrt{2} \quad (15c)$$

and

$$\text{erf}(x) = 1 - (a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5)e^{-x^2} \quad (15d)$$

where

$$t = \frac{1}{1 + px}$$

$$|\varepsilon(x)| < 1.5 \cdot 10^{-7}$$

where $p = 0.3275911$, $a_1 = 0.254829592$, $a_2 = -0.284496736$, $a_3 = 1.421413741$, $a_4 = -1.453152027$, $a_5 = 1.061405429$ and $\varepsilon(x)$ is the limit of approximation error for $0 < x < \infty$ [11].

METHOD

Experimental measurements

Four peak parameters are to be measured from the experimental chromatographic peak, t_p , h_p , A and M_1 . The peak retention time can be read directly from the experimental peak. For high accuracy in determining band broadening factors with the proposed method, t_p and M_1 values should be measured accurately. It is important to note that the determined σ and τ values can only be as accurate as the measured t_p and M_1 values. In the integration of the peak for peak area (A) and peak centroid (M_1) with eqns. 4 and 5, care should be taken so that baseline noise does not affect the accuracy of the A and M_1 values.

Derived equations

In the proposed method, an iterative search algorithm is used to determine the σ and τ value of the experimental peak from measured t_p , h_p , A and M_1 values. For this purpose, equations were derived from the basic characteristics of the EMG peak model. By eliminating t_r from eqns. 6 and 14, we obtain

$$Z_p = \frac{t_p - (M_1 - \tau)}{\sigma} - \frac{\sigma}{\tau} \quad (16)$$

also, from eqns. 12 and 14,

$$\gamma = Z_p + \frac{\sigma}{\tau} \quad (17)$$

and, from eqn. 11,

$$\sigma = \frac{A}{h_p \sqrt{2\pi}} \cdot \exp(-\gamma^2/2) \quad (18)$$

By grouping together the experimentally measured information from the peak-shape factors, we find the following property of the EMG peak model:

$$\frac{M_1 - t_p}{\left(\frac{A}{h_p \sqrt{2\pi}}\right)} = \left(\frac{\tau}{\sigma} - \gamma\right) \exp(-\gamma^2/2) \quad (19)$$

We define

$$\beta = \left(\frac{\tau}{\sigma} - \gamma \right) \exp(-\gamma^2/2) = \left(\frac{\tau}{\sigma} - \gamma \right) \Gamma \tag{20}$$

where

$$\Gamma = \text{peak-height factor} = \exp(-\gamma^2/2) \tag{21}$$

which gives the peak-height ratio of the exponentially modified peak *versus* the original Gaussian peak, and

$$\beta^* = \frac{M_1 - t_p}{\frac{A}{h_p \sqrt{2\pi}}} = \text{the experimental } \beta \text{ value} \tag{22}$$

The following equations were developed to simplify the τ/σ calculations with improved accuracy: for $Z_p < -4$,

$$\frac{\tau}{\sigma} \approx \frac{1}{|Z_p|} \left[1 + \sum_{m=1}^6 (-1)^m \frac{(2m-1) \cdots 3 \cdot 1}{Z_p^{2m}} \right] \tag{23a}$$

for $-4 < Z_p < 0$,

$$\frac{\tau}{\sigma} \approx \sqrt{\frac{\pi}{2}} \sum_{n=1}^5 a_n t^n \tag{23b}$$

for $Z_p > 0$,

$$\frac{\tau}{\sigma} \approx \sqrt{2\pi} \exp(Z_p^2/2) - \sqrt{\frac{\pi}{2}} \sum_{n=1}^5 a_n t^n \tag{23c}$$

where a_n and t^n have been defined in eqn. 15.

Table I gives a few examples of calculated τ/σ , γ , Γ , β and Sk values. The τ/σ and β values from Table I are also plotted as a function of Z_p in Fig. 2. The characteristic relationships between τ/σ , Γ and Sk with β for the EMG model are depicted in Figs. 3 and 4.

Note that the semi-logarithmic β *versus* Z_p plot is nearly linear in the β range of experimental interest. We have taken advantage of this linear relationship to develop a much simplified computer algorithm. The linear approximation to the β *versus* Z_p plot takes the mathematical form

$$Z_p = 1.2028 + 2.474 \log \beta \tag{24}$$

Simplified computer algorithm

With a measured β^* value from an experimental chromatographic peak, one can quickly use Fig. 2 or 3 to read off the τ/σ value for an experimental chromatographic

TABLE I

CHARACTERISTICS OF EXPONENTIALLY MODIFIED GAUSSIAN PEAK

Z_p	τ/σ	γ	Γ^a	β^b	Sk
-10	0.0990	0.0981	0.9952	0.0009	0.002
-7	0.1401	0.1375	0.9906	0.0026	0.005
-6	0.1624	0.1584	0.9875	0.0039	0.008
-5	0.1928	0.1870	0.9827	0.0060	0.014
-4	0.2366	0.2265	0.9747	0.0099	0.024
-3	0.3046	0.2829	0.9608	0.0208	0.049
-2.5	0.3544	0.3215	0.9496	0.0312	0.075
-2.0	0.4214	0.3732	0.9327	0.0450	0.117
-1.5	0.5158	0.4388	0.9082	0.0699	0.193
-1.0	0.6560	0.5244	0.8715	0.1147	0.330
-0.75	0.7526	0.5787	0.8458	0.1471	0.435
-0.50	0.8761	0.6414	0.8141	0.1911	0.572
-0.25	1.0381	0.7133	0.7754	0.2518	0.747
0	1.2533	0.7979	0.7274	0.3313	0.955
0.25	1.5487	0.8957	0.6696	0.4372	1.186
0.50	1.9639	1.0092	0.6010	0.5737	1.415
0.75	2.5687	1.1393	0.5226	0.7470	1.618
1.00	3.4774	1.2876	0.4365	0.9559	1.775
1.50	7.2047	1.6388	0.2611	1.4533	1.944
2.00	18.1011	2.0552	0.1210	1.9416	1.991
2.50	56.5477	2.5177	0.0320	2.2708	1.999
3.00	225.3349	3.0044	0.0110	2.4372	2.000

^a $\Gamma = \exp(-\gamma^2/2)$.

^b $\beta = \left(\frac{\tau}{\sigma} - \gamma\right) \exp(-\gamma^2/2)$.

peak by finding the point on the β axis that corresponds to the β^* value. From that point, the τ/σ value is read directly at the corresponding Z_p value at β^* . The corresponding Z_p and τ/σ values are determined digitally in the following way:

(1) Estimate Z_p :

(a) obtain the initial estimate $(Z_p)_0$ by substituting β^* for β in eqn. 24:

$$(Z_p)_0 = 1.2028 + 2.474 \log \beta^* \quad (25)$$

(b) update the Z_p value by using the following iterative calculations:

$$(\Delta Z_p)_n = 2.474 \log (\beta^*/\beta_n) \quad (26)$$

and

$$(Z_p)_{n+1} = (Z_p)_n + (\Delta Z_p)_n \quad (27)$$

(2) Calculate τ/σ from Z_p by using eqn. 23 or 13.

(3) Calculate γ from eqn. 17.

(4) Calculate β from eqn. 20.

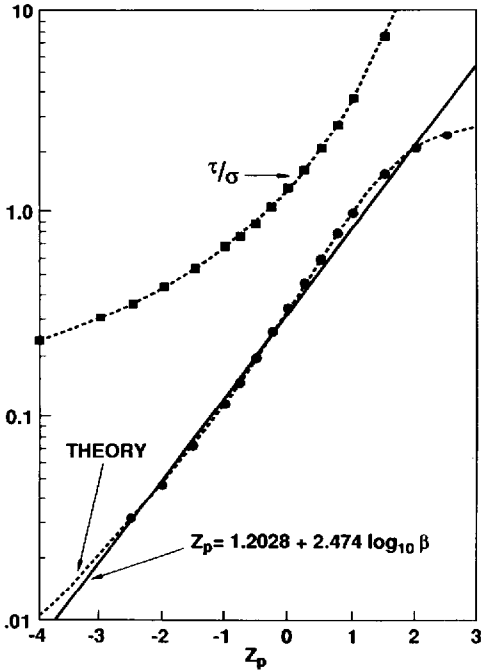


Fig. 2. Characteristics of exponentially modified Gaussian peak.

(5) Check the ΔZ_p value to terminate the iteration:

(a) if $|\Delta Z_p| > \delta$, go back to (1);

(b) if $|\Delta Z_p| < \delta$, end the iteration and continue to calculate σ from eqn. 18.

Proceed to calculate τ , peak skew, peak variance and other peak performance parameters described earlier. The value of δ is the search accuracy limit that can be adjusted to control the desired accuracy of the peak τ/σ value. We suggest a beginning

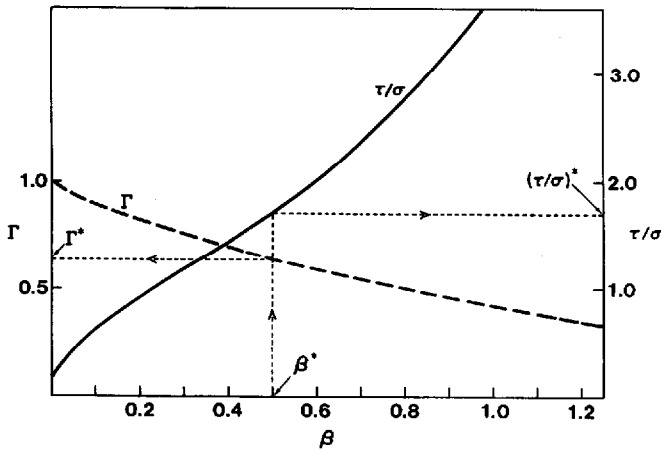


Fig. 3. Nomograph for extracting EMG peak functions.

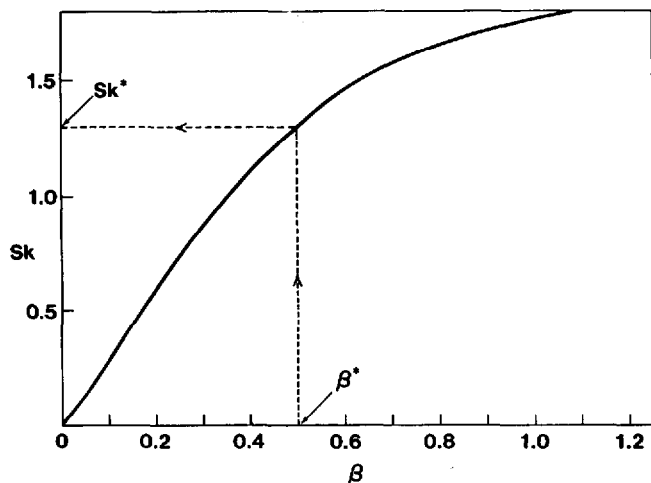


Fig. 4. Relationship between Sk and β value for EMG peak.

δ value of 0.02, as this produces a τ/σ value that is within about ± 0.01 of the actual value.

Computer program

The peak characterization method described above has been programmed for IBM-PC in BASIC and HP-1000 in FORTRAN-77. Special attention was paid to interpolate data points at the peak maximum so that an accurate determination of the experimental t_p value can be made. As was mentioned earlier, the accuracy of the present method is directly affected by how well t_p and M_1 values can be determined.

RESULTS AND DISCUSSION

The performance of the proposed method has been tested by a computer simulation study and by application to actual chromatographic data. The results of these studies will be presented in Part II [12].

REFERENCES

- 1 J. J. Kirkland, W. W. Yau, H. J. Stoklosa and C. H. Dilks Jr., *J. Chromatogr. Sci.*, 15 (1977) 303.
- 2 E. Grushka, *Anal. Chem.*, 44 (1972) 1733.
- 3 W. W. Yau, *Anal. Chem.*, 49 (1977) 395.
- 4 R. E. Pauls and L. B. Rogers, *Anal. Chem.*, 49 (1977) 625.
- 5 R. E. Pauls and L. B. Rogers, *Sep. Sci. Technol.*, 12 (1977) 395.
- 6 J. P. Foley and J. G. Dorsey, *Anal. Chem.*, 55 (1983) 730.
- 7 J. P. Foley and J. G. Dorsey, *J. Chromatogr. Sci.*, 22 (1984) 40.
- 8 J. V. H. Schudel and G. Guiochon, *J. Chromatogr.*, 457 (1988) 1.
- 9 J. C. Sternberg, *Adv. Chromatogr.*, 2 (1966) 205.
- 10 W. W. Yau, J. J. Kirkland and D. D. Bly, *Modern Size-Exclusion Liquid Chromatography*, Wiley, New York, 1979, Ch. 4 and 9.
- 11 M. Abramowitz and I. A. Stegun (Editors), *Handbook of Mathematical Functions*, Dover, New York, 1965, equation 7.1.26.
- 12 W. W. Yau, S. W. Rementer, J. M. Boyajian, K. B. Lim, J. J. DeStefano and J. J. Kirkland, *J. Chromatogr.*, in press.
- 13 M. S. Jeansonne and J. P. Foley, *J. Chromatogr. Sci.*, 29 (1991) 258.