Note

Least squares analysis of chromatographic tracings

D. R. COOPER*

Department of Chemistry, University of York, York YO1 5DD (Great Britain)

(Received February 26th, 1973)

In recent years, mathematical techniques for the resolution enhancement of overlapping chromatographic peaks have been proposed. The recommended procedures depend either upon a least squares fitting of the heights, shapes and positions of the individual peaks to the overall tracing[1–4], or upon the use of Fourier transform methods[5]. In addition, electronic devices have become available to permit the visual fitting of tracings by continuous adjustment of the contributions of the component peaks.

Although the simple method presented here is somewhat less versatile than certain of those referred to, it may nevertheless be valuable under circumstances in which resolution problems arise too infrequently, or are in general too straightforward, to justify the establishment of a more sophisticated system. It is particularly applicable to the repetitive analysis of mixtures, where the individual components are reasonably well characterised chromatographically, but cannot be perfectly separated on the available experimental set-up. It possesses the specific advantages that the positions and shapes of the component peaks do not enter explicitly into the basic mathematical treatment, and that the solution is analytic, rather than iterative. In consequence, relatively little computational effort is required.

Consider an experimental chromatographic tracing that is a composite of N individual peaks, of which the j'th alone would have a height h_j. Let the height of the tracing at the i'th of M positions along its length be H_i, to which the contribution of the j'th component is given by: $a_{i,j} \cdot h_j$. So the predicted height at the i'th position, S_i, is given by:

$$S_i = \sum_{k=1}^{k=N} a_{i,k} \cdot h_k$$

The M predicted heights must be fitted as closely as possible to the experimental heights H_i. This is achieved by minimising the sum of the squared deviations with respect to each h_j. Accordingly:

$$\partial/\partial h_j \sum_{i=1}^{i=M} (H_i - S_i)^2 = 0 \qquad \text{for each } j$$

---

* Present address: Department of Chemistry, University of Manchester, Manchester M13 9PL, Great Britain.

$$\therefore \sum_{i=1}^{i=M} 2(H_i - S_i) \frac{\partial S_i}{\partial h_j} = 0$$

Substituting for S_i:

$$\therefore \sum_{i=1}^{i=M} \left(H_i - \sum_{k=1}^{k=N} a_{i,k} \cdot h_k \right) a_{i,j} = 0$$

$$\therefore \sum_{i=1}^{i=M} a_{i,j} \sum_{k=1}^{k=N} a_{i,k} \cdot h_k = \sum_{i=1}^{i=M} H_i \cdot a_{i,j}$$

$$\therefore \sum_{i=1}^{i=M} \sum_{k=1}^{k=N} a_{i,j} \cdot a_{i,k} \cdot h_k = \sum_{i=1}^{i=M} H_i \cdot a_{i,j}$$

$$\therefore \sum_{k=1}^{k=N} \sum_{i=1}^{i=M} a_{i,j} \cdot a_{i,k} \cdot h_k = \sum_{i=1}^{i=M} H_i \cdot a_{i,j}$$

$$\therefore \sum_{k=1}^{k=N} h_k \sum_{i=1}^{i=M} a_{i,j} \cdot a_{i,k} = \sum_{i=1}^{i=M} H_i \cdot a_{i,j}$$

Putting:

$$\sum_{i=1}^{i=M} a_{i,j} \cdot a_{i,k} = A_{j,k}$$

and:

$$\sum_{i=1}^{i=M} H_i \cdot a_{i,j} = B_j$$

there results:

$$\sum_{k=1}^{k=N} A_{j,k} \cdot h_k = B_j$$

which is the j'th of a set of N simultaneous equations that define the optimal heights of the contributing peaks. This may be re-expressed in matrix notation as:

$$Ah = B$$

whence

$$h = A^{-1}B$$

Location of the component peaks may be by visual inspection, or by examination of the successive differentials of the trace[2,6]. The proportionality factors $a_{i,j}$ may then be found in whatever way is appropriate to the problem. Often, a formula of the Gaussian type will be applicable. For gel permeation chromatography, for instance:

$$a_{i,j} = \exp\left(\frac{-(V_i - V_j)^2}{2\sigma_j^2} \right)$$

where V_i and V_j are the elution volumes at the i'th data point taken from the trace,

and at the peak of the j'th component, respectively, and σ_j is the standard deviation of the peak about its mean position. In such a case, the area fraction of any one constituent may be readily found as:

$$h_j \cdot \sigma_j \bigg/ \sum_{j=1}^{j=N} h_j \cdot \sigma_j$$

h_j being the fitted height. Peaks of non-Gaussian shape may be described by means of tabulated data.

An equation similar to that derived above, but with the matrix elements $A_{j,k}$ and B_j represented by integrals, rather than summations, has previously been presented by Goldberg[4]. However, any associated refinement of the final fit must properly be viewed in the light of the additional computational labour involved in the numerical evaluation of these integrals.

A recent application of this technique has been to the analysis of the cyclic oligomers of certain polymers by means of gel permeation chromatography[7,8]. Fig. 1
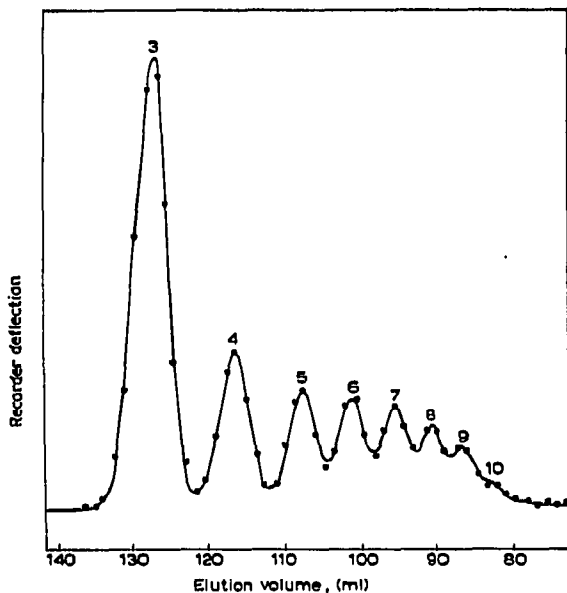

Fig. 1. Gel permeation chromatograph of the cyclic oligomers of poly(ethylene terephthalate). The fitted points are denoted by filled circles (●).

shows a typical experimental tracing obtained when cyclic oligomers extracted from poly(ethylene terephthalate) were separated by this method. (The identification of the cyclic components is according to the number of ethylene terephthalate repeat units per molecule.) The points fitted to the tracing by the described procedure are denoted as filled circles. In this instance, the peaks could be assumed Gaussian, and satisfactory information on their breadths was obtained from study of completely resolved materials eluting at similar positions.

ACKNOWLEDGEMENTS

REFERENCES

1 A. D. Booth, *Trans. Soc. Inst. Tech.*, 19 (1967) 12.
2 A. W. Westerberg, *Anal. Chem.*, 41 (1969) 1770.
3 A. H. Anderson, T. C. Gibb and A. B. Littlewood, *Anal. Chem.*, 42 (1970) 434.
4 B. Goldberg, *J. Chromatogr. Sci.*, 9 (1971) 287.
5 D. W. Kirmse and A. W. Westerberg, *Anal. Chem.*, 43 (1971) 1035.
6 J. R. Morrey, *Anal. Chem.*, 40 (1968) 905.
7 D. R. Cooper, *D. Phil. Thesis*, University of York, 1972.
8 D. R. Cooper and J. A. Semlyen, *Polymer*, in press.