

CHROM. 22 674

Direct deconvolution of Tung's integral equation using a multi-Gaussian function model for instrumental band broadening in gel-permeation chromatography

JING FENG* and XUNPEI FAN

Department of Applied Chemistry, Shanghai Jiao Tong University, 1954 Hua Shan Road, Shanghai 200030 (China)

(First received February 27th, 1990; revised manuscript received May 23rd, 1990)

ABSTRACT

A novel multi-Gaussian function model is reported for deconvoluting directly the Tung equation for instrumental band broadening (column dispersion) in gel-permeation chromatography. Using the Fourier transform technique, Tung's integral equation is reduced to a simple expression for the calculation of the true molecular weight distribution from an experimental chromatogram. An example of numerical calculation with computer simulation and two examples of experimental results for commercial polymers are given.

INTRODUCTION

Since the development of gel-permeation chromatography (GPC) by Moore [1], there has been increasing acceptance and use of this method for determining the molecular weight distributions (MWD) of resins and polymers, and better methods of processing the data and interpreting the results are needed. It has been shown by Tung and co-workers [2–4] and Hess and Kratz [5] that axial dispersion, which results in peak broadening, must be considered and corrected in order to obtain better agreement between calculated and experimental MWD results. For the difficult deconvolution of Tung's proposed equation, mathematical methods proposed by several workers [3–6] lead to some practical difficulties, such as oscillation of the dispersion-corrected results near the beginning and end of the chromatogram [3,4] and computational difficulties [5]. According to Tung's work [4], the Gaussian function is a reasonable approximation for many narrow-distribution chromatograms, and the chromatogram in GPC is considered as a multi-component model in this paper and each component is described by a Gaussian function. Hence the dispersion correction of Tung's equation is transformed into a non-linear regression analysis and a Fourier transform technique. The proposed method describes a polymer as a blend of components, each of them defined by its parameters and weight fraction. It provides the information needed in the development and engineering of a polymerization process and may be related to operating conditions.

MULTI-COMPONENT MODEL

A chromatogram in this study is assumed to be a blend of several components, each defined by its weight fraction and a set of parameters. A general equation for the normalized experimental chromatogram $f(v)$ is given by

$$f(v) = \sum_{i=1}^n C_i U_i(\bar{p}_i, v) \quad (1)$$

and

$$\sum_{i=1}^n C_i = 1, C_i > 0, i = 1, 2, \dots, n \quad (1a)$$

where n represents the number of components in the blend, C_i denotes the weight fraction of the i th component and U_i is the Gaussian equation describing the form of the i th component calculated for an eluent volume v , with parameters given by the vector \bar{p} :

$$U_i(\bar{p}_i, v) = \sqrt{h_i/\pi} \cdot \exp[-h_i(v - v_i)^2] \quad (2)$$

and

$$\bar{p}_i = (h_i, v_i), h_i < h, i = 1, 2, \dots, n \quad (2a)$$

h represents the dispersion factor. Hence $3n - 1$ parameters are needed to characterize a blend and a non-linear optimization procedure is employed.

DETERMINATION OF n , C_i AND VECTOR \bar{p}_i

One purpose of this work is to reproduce the whole chromatogram and the objective function of the non-linear regression is to minimize the sum of relative errors:

$$Q_n = \sum_{j=1}^m \left[f(v_j) - \sum_{i=1}^n C_i U_i(\bar{p}_i, v_j) \right]^2 \quad (3)$$

subject to

$$\sum_{i=1}^n C_i = 1, C_i > 0, h_i < h, i = 1, 2, \dots, n \quad (3a)$$

where m represents the number of experimental points. The unknown parameters n , C_i and \bar{p}_i should be found. The decision-making process involves assigning values to the parameters and we can seek optimum values of the parameters in the least-squares sense, that is, those values for which the sum of squares of the experimental deviations from the theoretical curve is minimized. This procedure is called least-squares fitting

and is a constrained non-linear optimization problem. Determination of the model starts with one component and the number of components is increased until an acceptable fit is obtained between the computed curve and experimental curve. Hence it can be understood that the procedure is binary optimization, one being the optimization of the value of n and the other the optimization of the values of C_i and \bar{p}_i for a given n . A flow chart of the procedure is illustrated in Fig. 1.

FOURIER TRANSFORM TECHNIQUE

The equation for dispersion correction proposed by Tung [2] is

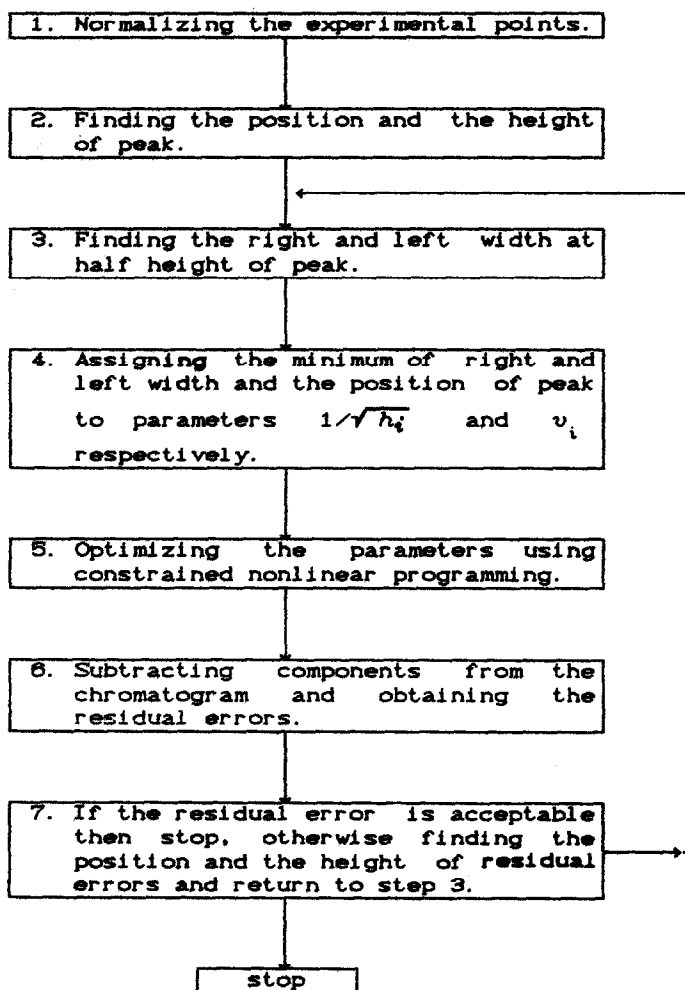


Fig. 1. Flow chart of non-linear regression for optimization of parameters.

$$f(v) = \int_{-\infty}^{+\infty} w(y)g(v-y)dy \quad (4)$$

where v and y are used to denote the eluent volume. This equation relates the observed chromatogram $f(v)$ to the distribution function $w(v)$ that would be obtained if axial dispersion effects were absent. The function $g(v-y)$ is the general expression of the dispersion function, which is an approximately Gaussian type [3]:

$$g(v) = \sqrt{h/\pi} \cdot \exp(-hv^2) \quad (5)$$

where h represents the dispersion factor. It is necessary to have some convenient methods of solving Tung's integral equation. The Fourier transform (FT) method [7] can be used to give a formal solution of eqn. 4. We can define the FT of $f(v)$, $w(v)$ and $g(v)$ by

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(v) \exp(ikv) dv \quad (6a)$$

$$W(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} w(v) \exp(ikv) dv \quad (6b)$$

$$G(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(v) \exp(ikv) dv \quad (6c)$$

We then take the FT of eqn. 4, applying the Faltung theorem [7] to the integral:

$$F(k) = \sqrt{2\pi} W(k) G(k) \quad (7a)$$

or

$$W(k) = \frac{F(k)}{\sqrt{2\pi} G(k)} \quad (7b)$$

Eqn. 6c can be solved as follows:

$$G(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sqrt{h/\pi} \cdot \exp(-hv^2 + ikv) dv \quad (8a)$$

Let

$$\xi = \sqrt{h}v - \frac{ik}{2\sqrt{h}} \text{ and } d\xi = \sqrt{h} dv$$

then

$$\begin{aligned} G(k) &= \frac{1}{\sqrt{2\pi}} \sqrt{h/\pi} \cdot \exp[-k^2/(4h)] \cdot \frac{1}{\sqrt{h}} \int_{-\infty}^{+\infty} \exp(-\xi^2) d\xi \\ &= \frac{1}{\sqrt{2\pi}} \cdot \exp[-k^2/(4h)] \end{aligned} \quad (8b)$$

Solving eqn. 6a for $F(k)$ combined with eqn. 5, by analogy with the method used above, the following expression is derived:

$$F(k) = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^n C_i \exp[-k^2/(4h_i) + ikv_i] \quad (9)$$

Combination of eqns. 7b, 8b and 9 gives

$$W(k) = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^n C_i \exp[-k^2/(4h_i) + ikv_i + k^2/(4h)] \quad (10)$$

Solving this equation for $W(k)$ and taking the inverse transform of the equation gives

$$\begin{aligned} W(v) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} W(k) \exp(-ikv) dk \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sum_{i=1}^n C_i \exp\{-[1/(4h_i) - 1/(4h)]k^2 - i(v - v_i)k\} dk \end{aligned}$$

For the validity of the above expression, it is reasonable to ask for h greater than h_i , and putting

$$\xi = \frac{1}{2} \sqrt{1/h_i - 1/h} \cdot k + \frac{i(v - v_i)}{\sqrt{1/h_i - 1/h}}$$

and

$$d\xi = \frac{1}{2} \sqrt{1/h_i - 1/h} dk$$

then

$$\begin{aligned}
 W(v) &= \frac{1}{2\pi} \sum_{i=1}^n C_i \exp[-(v - v_i)^2 / (1/h_i - 1/h)] \cdot \frac{2}{\sqrt{1/h_i - 1/h}} \int_{-\infty}^{+\infty} \exp(-\xi^2) d\xi \\
 &= \sum_{i=1}^n C_i \frac{1}{\sqrt{\pi} \sqrt{1/h_i - 1/h}} \cdot \exp[-(v - v_i)^2 / (1/h_i - 1/h)]
 \end{aligned}$$

Let

$$\rho_i = 1 / (1/h_i - 1/h)$$

then

$$W(v) = \sum_{i=1}^n C_i \sqrt{\rho_i / \pi} \cdot \exp[-\rho_i (v - v_i)^2] \quad (11)$$

This simple expression involves the parameter h , which is a function of the set of columns used, and h_i , v_i and C_i , which are related to the i th Gaussian function. The mathematical operations can be performed by using a computer. Several examples are given below.

RESULTS AND DISCUSSION

A numerical example is used to demonstrate the validity of the procedure. We assume that there is a known two-component MWD for $w(v)$, $C_1 = 0.65$, $\bar{p}_1 = (0.5, -1.2)$, $C_2 = 0.35$ and $\bar{p}_2 = (0.45, 1.5)$, and that the value of the dispersion factor h is taken as 0.8 over the entire range. The simulated chromatogram $f(v)$ shown in Fig. 2 (dashed line) is computed by direct integration of eqn. 1. The optimum values of non-linear regression are $C_1 = 0.651$, $\bar{p}_1 = (0.308, -1.19)$, $C_2 = 0.349$ and $\bar{p}_2 = (0.288, 1.50)$. The results of dispersion correction after Fourier transformation are shown in Fig. 2. It is obvious that the procedure works well and effectively.

The dashed line plotted in Fig. 3 represents the chromatogram of commercial polystyrene (PS) with a narrow MWD and the crosses indicate the results of optimization by non-linear regression for which the optimum values are given in Table I. The relative error is $1.89 \cdot 10^{-2}$ for one component and decreases to $1.44 \cdot 10^{-3}$ with two components and to $4.50 \cdot 10^{-4}$ for three components.

The solid line in Fig. 3 illustrates the chromatogram with a set of columns of greater resolving power.

The chromatogram for commercial poly(methyl methacrylate) (PMMA), which has a broader MWD, is shown in Fig. 4 (symbols as in Figs. 2 and 3), and the optimum values of the parameters are given in Table II. The relative error decreases from $4.48 \cdot 10^{-4}$ to $3.10 \cdot 10^{-4}$ when the number of components is increased from one to two. It is apparent that the effect of the dispersion correction for a broader MWD is not as great as that for a narrow MWD, for which the dispersion is important. Satisfactory

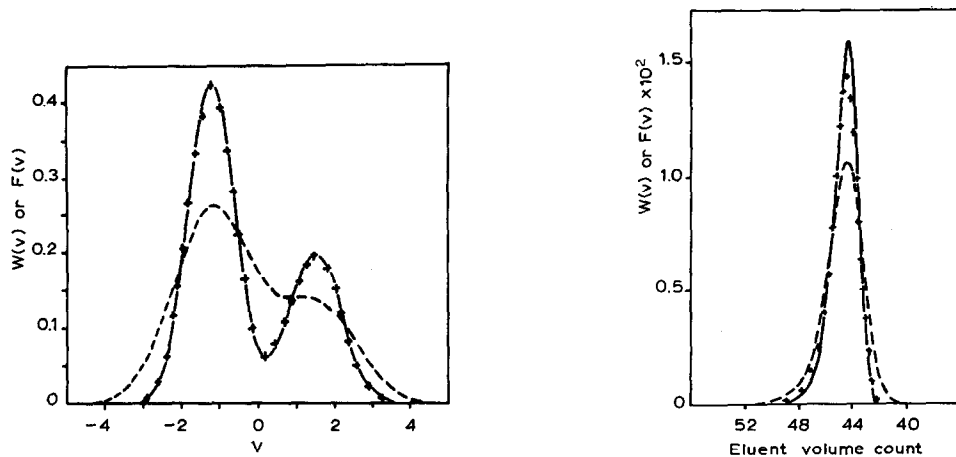


Fig. 2. Uncorrected and corrected chromatograms for a known MWD. Dashed line, uncorrected chromatogram; solid line, known chromatogram, x, corrected chromatogram.

Fig. 3. Uncorrected and corrected chromatograms for a narrow MWD polystyrene sample. Symbols as in Fig. 2.

TABLE I
OPTIMUM VALUES OF PARAMETERS FOR PS

| i | C_i | h_i | v_i | Q_i |
|-----|--------|-------|-------|----------------------|
| 1 | 0.84 | 0.398 | 44.08 | $1.89 \cdot 10^{-2}$ |
| 2 | 0.12 | 0.430 | 46.10 | $1.44 \cdot 10^{-3}$ |
| 3 | 0.0399 | 0.301 | 47.40 | $4.50 \cdot 10^{-4}$ |

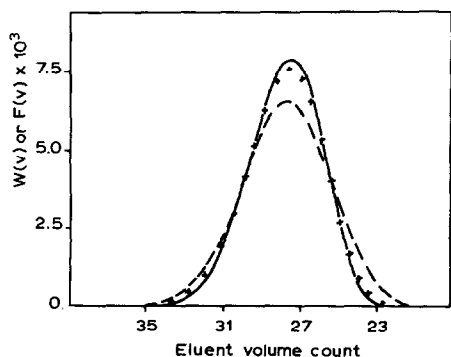


Fig. 4. Uncorrected and corrected chromatograms for a PMMA sample. Symbols as in Fig. 2.

TABLE II
OPTIMUM VALUES OF PARAMETERS FOR PMMA

| i | C_i | h_i | v_i | Q_i |
|-----|-------|-------|-------|----------------------|
| 1 | 0.99 | 0.092 | 28.25 | $4.48 \cdot 10^{-4}$ |
| 2 | 0.01 | 0.050 | 31.00 | $3.10 \cdot 10^{-4}$ |

agreement of the dispersion corrections is obviously obtained, but it should be noted that use of dispersion corrections, as shown in Figs. 3 and 4, is not a substitute for use of columns with the highest plate number available.

In the calculation of dispersion corrections for PS and PMMA in Figs. 3 and 4,

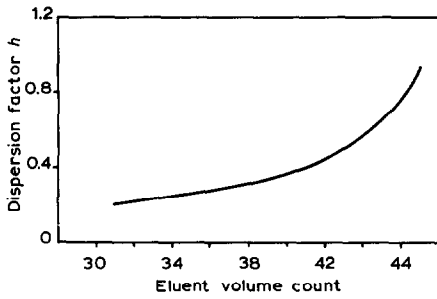


Fig. 5. Relationship between dispersion factor h and eluent volume v .

TABLE III
OPTIMUM VALUES OF PARAMETERS FOR PS

| i | C_i | h_i | v_i | Q_i |
|-----|-------|-------|-------|----------------------|
| 1 | 0.85 | 0.403 | 44.10 | $2.02 \cdot 10^{-2}$ |
| 2 | 0.13 | 0.410 | 46.17 | $1.14 \cdot 10^{-3}$ |
| 3 | 0.02 | 0.290 | 47.45 | $6.67 \cdot 10^{-4}$ |

TABLE IV
OPTIMUM VALUES OF PARAMETERS FOR PMMA

| i | C_i | h_i | v_i | Q_i |
|-----|-------|-------|-------|----------------------|
| 1 | 0.98 | 0.090 | 28.30 | $4.82 \cdot 10^{-4}$ |
| 2 | 0.02 | 0.052 | 31.00 | $2.28 \cdot 10^{-4}$ |

a variable value of h was used. The papers by Tung and co-workers [2-4] and Hess and Kratz [5] indicate that h varies, being smaller for smaller eluent volumes. The value of h was determined for monodisperse PS standards by using a Gaussian peak distribution as in eqn. 5. Fig. 5 shows the relationship between h and the eluent volume v .

The two examples shown above were tested in another experiment (chromatograms not shown), and the results of optimum programming are given in Tables III and IV.

We consider that the majority of residual relative errors are produced by the assumption that a Gaussian distribution can be used to determine the dispersion factor h . As mentioned above, the monodisperse PS samples are standard substances for the determination of h but the MWDs of these samples are like the narrow MWD shown in Fig. 3. The assumption leads to the method proposed here, not for total deconvolution. We can deduce from Tables III and IV that the values of the parameters are reproducible experimentally. The uniqueness of the parameters is ensured by the algorithm with optimization theory. Although the parameters may be variable in different optimum algorithms, the final $w(v)$ results are not affected. A more detailed calculation may be needed to test the method described here.

REFERENCES

- 1 J. C. Moore, *J. Polym. Sci., Part A*, 2 (1964) 835.
- 2 L. H. Tung, J. C. Moore and G. W. Knight, *J. Appl. Polym. Sci.*, 10 (1966) 917.
- 3 L. H. Tung, *J. Appl. Polym. Sci.*, 10 (1966) 375.
- 4 L. H. Tung, *J. Appl. Polym. Sci.*, 10 (1966) 1271.
- 5 M. Hess and R. F. Kratz, *J. Polym. Sci., Part A-2*, 4 (1966) 731.
- 6 W. N. Smith, *J. Appl. Polym. Sci.*, 11 (1967) 639.
- 7 P. M. Morse and H. Feshback, *Methods of Theoretical Physics*, McGraw-Hill, New York, 1953, Part I, Sect. 8.4, p. 464.