JOURNAL OF APPLIED POLYMER SCIENCE VOL. 10, PP. 375–385 (1966)

Method of Calculating Molecular Weight Distribution Function from Gel Permeation Chromatograms

L. H. TUNG, *Physical Research Laboratory, The Dow Chemical Company, Midland, Michigan*

Synopsis

An integral equation taking account of the limited resolution of the chromatographic columns is given to relate the gel permeation chromatogram and the true molecular weight distribution function. Three approaches to solve the integral equation are described. The first approach provides a special solution for the log-normal molecular weight distribution function; the other two approaches give two numerical solutions for general distribution functions. The use of these solutions in the treatment of gel permeation chromatography data is discussed.

Introduction

Gel permeation chromatography since its introduction by Moore[1] has been widely accepted as a convenient method of determining the molecular weight distribution of polymeric compounds. The method is rapid, and has good reproducibility and accuracy. However, its resolution, like those of other chromatographic techniques, is not unlimited. The chromatogram of a monomeric compound appears not as a straight line but as a bell-shaped Gaussian curve, as shown in Figure 1. The position of the peak of the Gaussian curve depends on the molecular weight of the compound; the area under the curve is proportional to the weight or concentration of the compound in solution, and the width of the Gaussian curve depends on the resolution of the chromatographic columns. For a polydispersed sample the chromatogram is a composite of the Gaussian curves of all its components. The total area under the curve is still proportional to the concentration of the sample but the height of the curve does not reflect the relative abundance of the components at the corresponding eluent volumes, as it depends also on the abundance of the neighboring components. At the two extremities of the chromatograms there are curve portions representing components which do not even exist in the sample. It is clear that this overlapping and diffused pattern of the chromatogram must be accounted for in the calculation of the true molecular weight distribution. The present paper describes a method to treat this problem.

Formulation of the Problem

Because the detailed mechanism in a gel chromatographic column is not yet fully understood, we shall treat the problem from a phenomenological point of view.

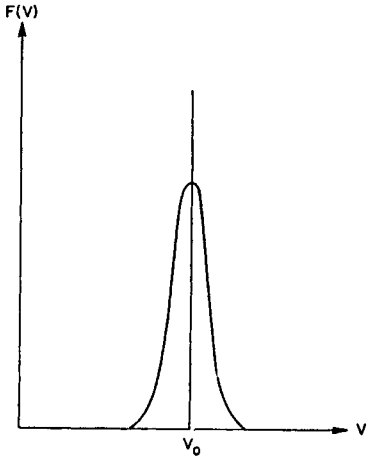

Fig. 1. Chromatogram of a monomeric compound.

It has been demonstrated that in gel permeation chromatography the eluent volume is proportional to the logarithm of the molecular weight. For such relationship we write

$$v = C_1 - C_2 \ln M \tag{1}$$

We shall use v and y interchangeably to denote the eluent volume; y is used mainly to denote the eluent volume as the variable under the definite integral sign. We shall use $W(y)$ or $W(v)$ to denote the distribution function with respect to the eluent volume. $W(y)$ is equivalent also to the chromatogram which would be obtained if the resolution were infinitely large. The differential molecular weight distribution function, $w(M)$, is then

$$w(M) = C_2 N W(y)/M \tag{2}$$

where N is the normalization factor for $W(y)$

$$N = 1 \Big/ \int_{-\infty}^{\infty} W(y)dy \tag{3}$$

We shall use $F(v)$ to denote the chromatogram; $F(v)$, like $W(y)$, is not normalized.

The Gaussian shape chromatogram shown in Figure 1 can be represented by the equation

$$F(v) = A\sqrt{h/\pi}\, e^{-h(v - v_0)^2} \tag{4}$$

where v_0 is the eluent volume at the peak of the curve, A is a constant related to the concentration of the compound; and h is the resolution factor. For infinitely large resolution h approaches infinity and the curve reduces to a straight line. For a multicomponent system

$$F(v) = \sum_i A_i \sqrt{h_i/\pi}\, e^{-h_i(v - v_{0i})^2} \tag{5}$$

If the number of components is large and a distribution function $W(y)$ can be used to denote the abundance of the components in the system then

$$F(v) = \int_{v_a}^{v_b} W(y) \sqrt{h/\pi}\, e^{-h(v-y)^2} dy \tag{6}$$

where v_a is the initial eluent volume and v_b the final eluent volume of the chromatogram. Equation (6) is similar to transform equations, and no general solution for $W(y)$ is available.

Special Solution for the Log-Normal Distribution

A familiar molecular weight distribution function is the log-normal distribution

$$w(M) = [1/(\beta\sqrt{\pi}M)] \exp\{-(1/\beta^2) \ln^2 (M/M_0)\} \tag{7}$$

Transformation to the y coordinate gives

$$W(y) = [C_2/(N\beta\sqrt{\pi})] \exp\{-[1/(C_2^2\beta^2)](y - y_0)^2\} \tag{8}$$

where $y_0 = C_1 - C_i \ln M_0$. Equation (8) is Gaussian. Hence β is a parameter denoting the breadth of the distribution and y_0 is the eluent volume at the peak of the $W(y)$ curve. Equation (6) can be solved readily for such a case if h is a constant. Substituting eq. (8) for $W(y)$ in eq. (6) and integrating, we have

$$F(v) = \{C_2\sqrt{h}/[N\sqrt{\pi(1 + hC_2^2\beta^2)}]\} \times \exp\{-[h/(1 + C_2^2\beta^2 h)](v - y_0)^2\} \tag{9}$$

We see that the chromatogram for such distribution retains the Gaussian form. The peak of $F(v)$ remains at the same position as $W(y)$. The width of $F(v)$ is broader than $W(y)$ except when h approaches infinity. In the latter case $F(v)$ reduces to $W(y)$. Through eq. (9), β and M_0 can be derived from the chromatogram, and $w(M)$ is hence determined.

Numerical Solution for General Distributions by Using the Gaussian Quadrature Formula

The Gaussian quadrature approximation for the integral equation, eq. (6), can be written as

$$F(v) = [(v_b - v_a)/2] \sum_{i=1}^{n} G_i W(y_i) \sqrt{h_i/\pi}\, e^{-h_i(v-y_i)^2} \tag{10}$$

where

$$y_i = [(v_b - v_a)/2]x_i + [(v_b + v_a)/2]$$

The weight coefficients G_i and abscissas x_i for Gaussian quadrature have been tabulated.[2] The larger the number of subdivisions n for the interval (v_a, v_b) the more accurate is the approximation. We have tested this

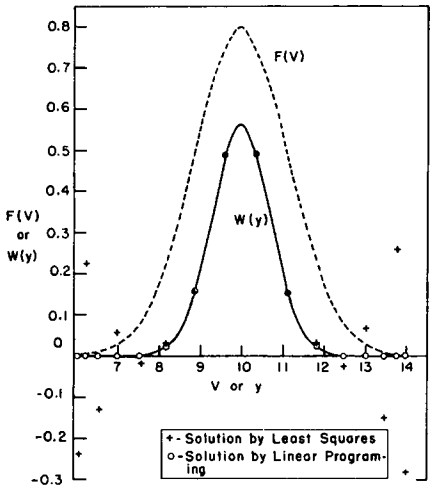

Fig. 2. Comparison of the solutions by Gaussian quadrature method with a known distribution function (Gaussian).

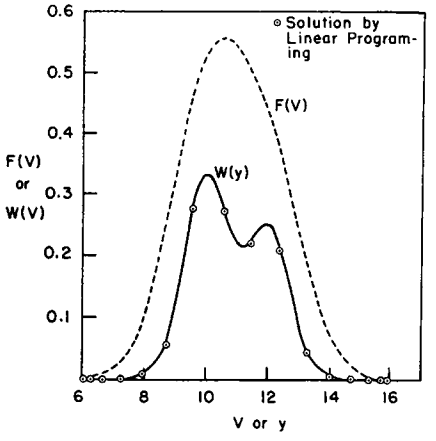

Fig. 3. Comparison of the solution by Gaussian quadrature method with a known binodal distribution function.

approximation with known hypothetical functions for $F(v)$ and $W(y)$ and found that for $n = 16$ the results agree within 1% to the values of $F(v)$ obtained by direct integration.

In eq. (10) the function $W(y)$ becomes n discrete unknowns $W(y_i)$. Thus if n pairs of v and $F(v)$ values are selected from the experimental chromatogram, the n $W(y_i)$ can be solved from the set of n simultaneous linear algebraic equations. In carrying out such a procedure, we have found, however, that the solutions for $W(y_i)$ near v_a and v_b become violently oscillatory about the y axis. Such oscillation is dampened only slightly if more than n points are selected from the chromatogram and $W(y_i)$ are solved by the method of least squares. This difficulty finally is overcome

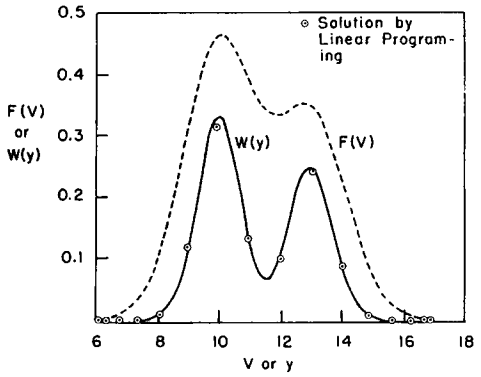

Fig. 4. Comparison of the solution by Gaussian quadrature method with a known binodal distribution function.

by using linear programming on a digital computer whereby all negative values of $W(y_i)$ are rejected as solutions.

To demonstrate these results we used a $F(v)$, with an available analytical solution of $W(y)$. The value of h is taken as 0.6 over the entire range. These functions are shown by the curves in Figure 2. The crosses in the figure indicate points calculated by the method of least squares and the circles indicate points calculated by linear programming.

Figure 2 shows that only half of the 16 points fall on the main part of $W(y)$ curve. For accurate description of $W(y)$ more points (larger n) are required. This is further illustrated in Figures 3 and 4, where more complex chromatograms are used as examples. Points shown in both figures were computed by linear programming.

The normalization factor and the average molecular weights are:

$$N = 1\Big/\left[\frac{v_b - v_a}{2}\sum_{i=1}^{n} G_i W(y_i)\right] \tag{11}$$

$$\bar{M}_n = 1\Big/\left[N\frac{v_b - v_a}{2}\sum_{i=1}^{n} G_i W(y_i)/M_i\right] \tag{12}$$

$$\bar{M}_w = N\frac{v_b - v_a}{2}\sum_{i=1}^{n} G_i W(y_i) M_i \tag{13}$$

$$\bar{M}_z = \left[\sum_{i=1}^{n} G_i W(y_i) M_i^2\right]\Big/\left[\sum_{i=1}^{n} G_i W(y_i) M_i\right] \tag{14}$$

where M_i is the molecular weight corresponding to the eluent volume y_i.

Numerical Solution for General Distributions by Using Polynomial Expansion

The distribution function $W(y)$ has the characteristics of being continuous and approaching zero as y approaches $\pm\ \infty$. Such a function can be represented by a polynomial of the form:

$$W(y) = e^{-p^2(y - y_0)^2}\sum_{=0}^{n} R_i(y - y_0)^i \tag{15}$$

where R_i are the coefficients of the polynomial and p and y_0 are two additional adjustable parameters. The number of terms required for an accurate description of $W(y)$ depends on the choices of p and y_0 and the complexity of the function itself [eq. (15) reduces to the Gaussian equation when $n = 0$]. If h is a constant and $W(y)$ is represented by the above polynomial, eq. (6) can be integrated. As neither $W(y)$ nor $F(v)$ is normalized, we may eliminate the factor $\sqrt{h/\pi}$ from eq. (6) and write the integral equation for constant h as

$$F(v) = \int_{-\infty}^{+\infty} W(y)e^{-h(v-y)^2}dy \tag{16}$$

The limits of integration have been extended to $\pm\infty$. Substituting eq. (15) into eq. (16) and performing the integration, we have

$$F(v) = e^{-(p^2h/l^2)(v-y_0)^2} \sum_{i=0}^{n} (h/l^2)^i (v-y_0)^i \sum_{j=i}^{n} R_{jj} C_2 Q_{j-i} \tag{17}$$

where ${}_jC_i$ are binomial coefficients, and

$$l = \sqrt{p^2+h}$$

$$Q_0 = \sqrt{\pi}/l$$

$$Q_i = [\Gamma(1/2 + 1/2)]/l^{i+1}$$

$$= \frac{1 \times 3 \times 5 \times \ldots (i-1)\sqrt{\pi}}{2^{i/2} l^{i+1}} \qquad \text{For even } i$$

$$Q_i = 0 \qquad \text{For odd } i$$

$$Q_{j-i} = \frac{1 \times 3 \times 5 \times \ldots (j-i-1)\sqrt{\pi}}{2^{(j-i)/2} l^{j-i+1}} \qquad \text{For even } (j-i)$$

We can rewrite eq. (17) as

$$F(v) = e^{-q^2(v-y_0)^2} \sum_{i=0}^{n} U_i (v-y_0)^i \tag{18}$$

where $q = p\sqrt{h}/l$

$$U_i = (h/l^2)^i \sum_{j=i}^{n} R_{jj} C_i Q_{j-i}$$

The problem reduces now to the fitting of eq. (18) to experimental chromatograms. Then from the parameters q and y_0 and the coefficients U_i the corresponding parameters and coefficients for $W(y)$ can be calculated. The explicit forms for R_i are

$$R_n = U_n (l^2/h)^n / Q_0 \tag{19}$$

and

$$R_{n-i} = \frac{U_{n-i}(l^2/h)^{n-i} \sum_{j=0}^{(i-1)} R_{n-j\,n-j} C_{n-i}\, Q_{i-j}}{{}_{n-i}C_{n-i}\, Q_0} \tag{20}$$

The fitting of eq. (18) to experimental chromatograms can be conveniently done by the use of Hermite polynomials.

We replace $q(v - y_0)$ in eq. (18) by z and write

$$\bar{F}(z) = e^{z^2} \sum_{i=0} b_i H_i(z) \tag{21}$$

where $H_i(z)$ are Hermite polynomials of degree i

$$H_i(z) = \sum_{j=0}^{m} (2z)^{i-2j} \frac{i!}{(i-2j)!\,j!} (-1)^j \tag{22}$$

and

$$m = i/2 \text{ when } i \text{ is even}$$

$$m = (i-1)/2 \text{ when } i \text{ is odd}$$

A property of Hermite polynomials is

$$\int_{-\infty}^{\infty} e^{-z^2} H_i(z) \cdot H_j(z) dz = 2^i i! \sqrt{\pi} \delta_{i,j} \tag{23}$$

Thus the coefficients b_i can be expressed by

$$b_i = \left[\int_{-\infty}^{\infty} \bar{F}(z) H_i(z) dz \right] \Big/ 2^i i! \sqrt{\pi} \tag{24}$$

Let μ_i denote the ith moment of $\bar{F}(z)$ or

$$\mu_i = \int_{-\infty}^{\infty} \bar{F}(z) z^i dz \tag{25}$$

Then

$$b_i = \left\{ \sum_{j=0}^{m} 2^{i-2j} \mu_{i-2j} \frac{i!}{(i-2j)!\,j!} (-1)^j \right\} \Big/ 2^i i! \sqrt{\pi} \tag{26}$$

and

$$\bar{F}(z) = \frac{e^{-z^2}}{\sqrt{\pi}} \sum_{i=0}^{n} \frac{1}{2^i i!} \left[\sum_{j=0}^{m} 2^{i-2j} \mu_{i-2j} \frac{i!(-1)^j}{(i-2j)!\,j!} \right] \sum_{j=0}^{m} (2z)^{i-2j} \frac{i!(-1)^j}{(i-2j)!\,j!} \tag{27}$$

By comparing coefficients of like power of $(v - y_0)$ in eqs. (18) and (27) we have:

For even n and even i,

$$U_i = \frac{q^i}{\sqrt{\pi}} \sum_{k=i/2}^{n/2} \frac{2^{i-2k}(-1)^{k-i/2}}{i!(k-i/2)!} \sum_{j=0}^{k} 2^{2k-2j}\mu_{2k-2j}\frac{(2k)!(-1)^j}{(2k-2j)!j!} \tag{28}$$

For even n and odd i,

$$= \frac{q^i}{\sqrt{\pi}} \sum_{k=(i-1)/2}^{(n/2)-1} \frac{2^{i-2k-1}(-1)^{k-(i-1)/2}}{i![k-(i-1)/2]!} \times \sum_{j=0}^{k} 2^{2k+1-2j}\mu_{2k+1-2j}\frac{(2k+1)!(-1)^j}{(2k+1-2j)!j!} \tag{29}$$

For odd n the upper limits of the first summation sign in both eqs. (28) and (29) are replaced by $(n-1)/2$.

At this point we cannot yet calculate the moments μ_i as they contain two undetermined parameters, q and y_0. These two parameters are solved by arbitrarily putting the coefficients b_1 and b_2 to zero. Such a step is equivalent to making the higher terms in eq. (21) the difference between $F(v)$ and a Gaussian curve having the same first and second moments about y_0. Equation (21) is expected to converge faster than with other values for q and y_0.

The conditions for making b_1 and b_2 zero are

$$\mu_1 = 0 \tag{30}$$

$$2\mu_2 - \mu_0 = 0 \tag{31}$$

Let μ_1' denote the moments of $F(v)$ with respect to v, i.e.,

$$\mu_1' = \int_{-\infty}^{\infty} F(v)v^i dv \tag{32}$$

We find then

$$\mu_0 = q\mu_0' \tag{33}$$

$$\mu_1 = (\mu_1' - y_0\mu_0')q^2 \tag{34}$$

$$\mu_2 = (\mu_2' - 2y_0\mu_1' + y_0^2\mu_0')q^3 \tag{35}$$

Substituting eqs. (33), (34), and (35) into eqs. (30) and (31) we obtain

$$y_0 = \mu_1'/\mu_0' \tag{36}$$

$$q = \mu_0'/\sqrt{2(\mu_2'\mu_0' - \mu_1'^2)} \tag{37}$$

The moments μ_0', μ_1', and μ_2' can be determined numerically from the chromatogram. The Gaussian quadrature formula is again most useful in obtaining these moments. Once y_0 and q are determined, all the moments μ_i can be evaluated from the chromatogram.

Figure 5 shows the fit of eq. (18) to the $F(v)$ curve used in Figure 4. The curve corresponding to $n = 16$ (16 terms) is indistinguishable from the original $F(v)$. The comparison of $W(y)$ computed by using the 16-term polynomial and $W(y)$ obtained analytically is shown in Figure 6.

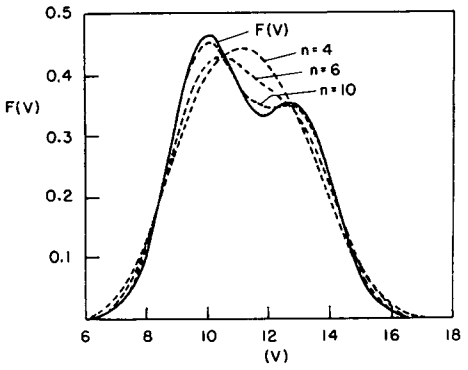

Fig. 5. Curve fitting by Hermite polynomials.

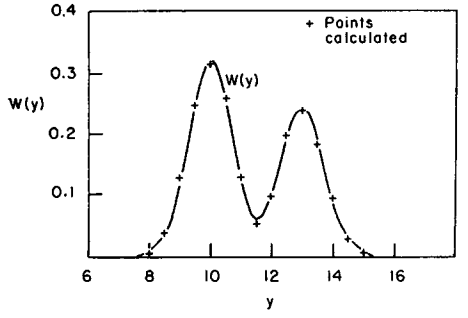

Fig. 6. Comparison of the solution by the polynomial expansion method with a known binodal distribution function.

The normalization factor and various average molecular weights are:

$$N = 1 \Big/ \left[\sum_{i=0}^{n} R_i Q_i' \right] \tag{38}$$

$$\bar{M}_n = M_0 \Big/ \left[N e^{1/(4C_2^2p^2)} \sum_{i=0}^{n} R_i \sum_{j=0}^{i} {}_iC_j \, Q'_{i-j} [1/(2C_2^2p^2)]^j \right] \tag{39}$$

$$\bar{M}_w = M_0 N e^{1/(4C_2^2p^2)} \sum_{i=0}^{n} R_i \sum_{j=0}^{i} {}_iC_j \, Q'_{i-j} [-1/(2C_2^2p^2)]^j \tag{40}$$

$$\bar{M}_z = M_0 e^{3/(4C_2^2p^2)} \frac{\sum_{i=0}^{n} R_i \sum_{j=0}^{i} {}_iC_j \, Q'_{i-j} [-1/(C_2^2p^2)]^j}{\sum_{i=0}^{n} R_i \sum_{j=0}^{i} {}_iC_j \, Q'_{i-j} [-1/(2C_2^2p^2)]^j} \tag{41}$$

where

$$Q_i' = (l/p)^{i+1} Q_i$$

and

$$M_0 = e^{(c_1 - y)/c_2}$$

Discussion

The special solution for the log-normal distribution should be of little value in practice. Though the log-normal distribution function has been frequently used to correlate fractionation data, the fit has seldom been very good. It was used mainly because of the lack of precision in conventional fractionation technique. It seems unwise to gain some simplicity in calculation at the expense of sacrificing the precision of the gel permeation chromatography.

The two numerical solutions for general distributions are of equal precision and neither is practical without the use of a computer. A preference will probably evolve after they are tested through repeated uses. In both methods $W(y)$ is solved without using the calibration equation, eq. (1). The Gaussian quadrature method has the additional flexibility that a variable h with respect to v can be handled easily. The polynomial method requires a constant h: if h is not constant, an average h must be used as an approximation. Thus, if the resolution of the chromatographic columns varies with molecular weight, the Gaussian quadrature method is definitely preferred. On the other hand the computer program for the polynomial method is shorter and simpler. If one has the patience it is not impossible to carry out the calculations on a desk calculator. Fewer data points are also required to represent the chromatogram in the polynomial method. It will probably be preferred if h is relatively constant with respect to molecular weight.

In the case that the calibration relation does not obey eq. (1), eq. (2) must be modified. Equations (39), (40), and (41) for the calculation of average molecular weights are also invalid.

In order to make use of these solutions we now need to calibrate the columns not only for molecular weight but also for resolution. Standard samples of known molecular weights and known molecular weight distributions are therefore required.

The present calculations have the advantage of yielding the true molecular weight distribution with relatively few arbitrary assumptions or approximations. In addition, chromatograms of low resolution can be analyzed with equal efficiency. Accurate results are thus possible with shorter columns and reduced elution time.

The author is indebted to M. Klein of the Computations Research Laboratory, The Dow Chemical Company, for providing the solutions by linear programming.

References

1. Moore, J. C., *J. Polymer Sci.*, **A2**, 835 (1964).
2. Abramowitz, M., and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D. C., 1964.

Résumé

Une équation intégrale tenant compte de la résolution limitée des colonnes chromatographiques est donnée en vue de reliér les chromatogrammes par perméation sur gel et la fonction de distribution de poids moléculaires vrais. Trois approches pour solutionner l'équation intégrale sont décrites. La première approche apporte une spéciale pour la fonction de distribution du poids moléculaire normal; les deux autres approches fournissent deux solutions numériques pour les fonctions de distribution générale. L'utilisation de ces solutions dans le traitement des données de chromatographie par perméation sur gel est soumise à discussion.

Zusammenfassung

Eine Integralgleichung, welche die begrenzte Auflösung einer chromatographischen Säule berücksichtigt, wird zur Aufstellung einer Beziehung zwischen dem Gelpermeationschromatogramm und der wahren Molekulargewichtsverteilung angegeben. Drei Lösungen für die Integralgleichung werden beschrieben. Die erst ist eine spezielle Lösung für die log-normale Molekulargewichtsverteilungsfunktion. Die beiden anderen sind zwei numerische Lösungen für allgemeine Verteilungsfunktionen. Die Verwendung dieser Lösungen bei der Behandlung von Gelpermeationschromatographiedaten wird diskutiert.

Received October 12, 1965
Prod. No. 1303