JOURNAL OF APPLIED POLYMER SCIENCE VOL. 10, PP. 1271–1283 (1966)

# Method of Calculating Molecular Weight Distribution Function from Gel Permeation Chromatograms. III. Application of the Method

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

## Synopsis

Six examples were used to illustrate the application of a previously reported method of calculating molecular weight distribution from gel permeation chromatograms. These examples show that the correction for the imperfect resolution of the GPC column is important when the distribution is narrow but minor when the distribution is broad. They show also that the variation of the resolution factor $h$, defined previously, with eluent volume can be neglected in the calculation. When the chromatograms are very narrow in distribution or when they are complex some difficulties are encountered. The two computer programs written to implement the previously described numerical calculations are shown to be adequate for these difficult cases but there are also limitations.

## INTRODUCTION

In the determination of molecular weight distribution by gel permeation chromatography (GPC), one is interested in not only the positions but also the shapes of the peaks in the chromatogram. The shape of a peak is, however, influenced by the resolution of the GPC column. To correct for this effect Hess and Kratz[1] used an unsymmetrical function to represent the broadening of the peak caused by the imperfect resolution. In their treatment the distribution was assumed to consist of an arbitrary number of discrete species. Another proposal for this correction,. described in an earlier paper[2] of this series, used a symmetrical Gaussian function to represent the broadening effect and treated the distribution as a continuous function.

It has been demonstrated in a second paper[3] of this series that the computation scheme using Gaussian function and continuous distribution agreed well with the performance of a GPC column. Two computer programs have been written for this computation scheme. In this paper six examples are used to show how these computer programs can be advantageously applied to actual GPC chromatograms and to what extent chromatograms are affected by the imperfect resolution of the columns. Some limitations of the computation scheme are also discussed.

## COMPUTER PROGRAMS

The integral equation relating the corrected chromatogram, $W(y)$, to the experimental chromatogram, $F(v)$, is given by

$$F(v) = \int_{v_b}^{v_a} W(y) \sqrt{h/\pi}\, e^{-h(v-y)^2}\, dy \tag{1}$$

where $h$ is the resolution factor representing the degree of the broadening effect as defined in the earlier paper;[2] $v$ and $y$ are both eluent volumes, $y$ denoting the eluent volume as the variable under the definite integral sign; $v_a$ and $v_b$ are the initial and final eluent volume of the experimental chromatogram, respectively. The two computer programs described below were written to implement the two numerical methods of solving eq. (1) given in the earlier paper.[2]

### Program I: Solution by Using The Gaussian Quadrature Formula

The Gaussian quadrature approximation for eq. (1) can be written as

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

where

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

$G_i$ are the Gaussian weighing coefficients and $x_i$ are the Gaussian abscissas. In eq. (2) the function $W(y)$ becomes $n$ discrete unknowns, $W(y_i)$. Each pair of $F(v)$ and $v$ read from the chromatogram gives one equation of the above form. The best values of $W(y_i)$ are determined from $n$ or more of these equations. This optimization of $W(y_i)$ is done by linear programming in which only positive values of $W(y_i)$ are accepted. The program allows $n$ to be chosen from four different values: 16, 20, 24, and 32. The number of data points to be read from the chromatogram must be equal or greater than $n$. Six significant digits are used for the largest coefficient $(G_i\sqrt{h_i/\pi}\, e^{-h_i(v-y_i)^2})$ for $W(y_i)$ in the set of simultaneous equations. Any coefficients smaller than $10^{-6}$ times this largest coefficient are treated as zero.

### Program II: Solution by Using Polynomial Expansion

In this method the resolution factor $h$ is treated as a constant and eq. (1) is simplified to

$$F(v) = \int_{-\infty}^{+\infty} W(y)\, e^{-h(v-y)^2}\, dy \tag{3}$$

$F(v)$ is then replaced by a polynomial

$$F(v) = e^{-q^2(v-y_0)^2} \sum_{i=0}^{n'} U_i(v - y_0)^i \tag{4}$$

where $q$, $y_0$, and $U_i$ are adjustable parameters and coefficients. The function $W(y)$ is expanded similarly and term by term integration is carried out. Finally by the comparison of coefficients of like power of $(v - y_0)$, $W(y)$ is solved.

The success of this method depends greatly on how well the chromatogram can be fitted by eq. (4) with a finite number of terms. In this program this fitting is done by the method of moments making use of the Hermite polynomials. The number of terms, $n'$, is limited from 1 to 32 as $U_i$ for $i$ greater than 30 are usually smaller than $1 \times 10^{-47}$, the smallest nonzero number which can be handled by the digital computer used for the present calculations. By virtue of the above method of moments, when $n'$ is 1 or 2, eq. (4) reduces to the Gaussian function (the case $n' = 0$).

## EXAMPLES

### Example I: Single Peak, Broad Distribution

For this example the relation of the resolution factor $h$ with the eluent volume is shown by curve $A$ in Figure 1; the molecular weight–eluent volume relationship is given by

$$v = 211 - 21(\log M) \tag{5}$$

The GPC column for this and all other five examples consisted of three 4-ft. sections and had a plate count of 680/ft. (by method described in manual of the instrument). As the chromatogram of this example covered a range where $h$ varied about six-fold, program I was used first. The circles in Figure 2 show the results calculated with $n = 20$. The ordinate, $W^*(\log M)$, in the figure is the distribution function with respect to the logarithm of molecular weight. The points at both ends of the distribution curve oscillated and obviously do not represent the true distribution.

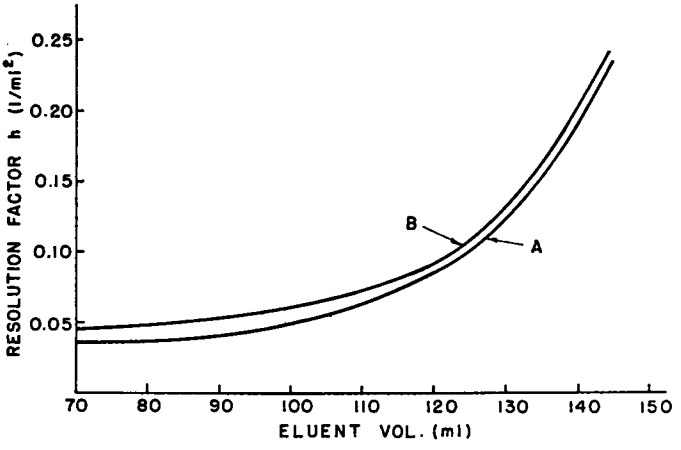


Fig. 1. Variation of the resolution factor $h$ with respect to the eluent volume.

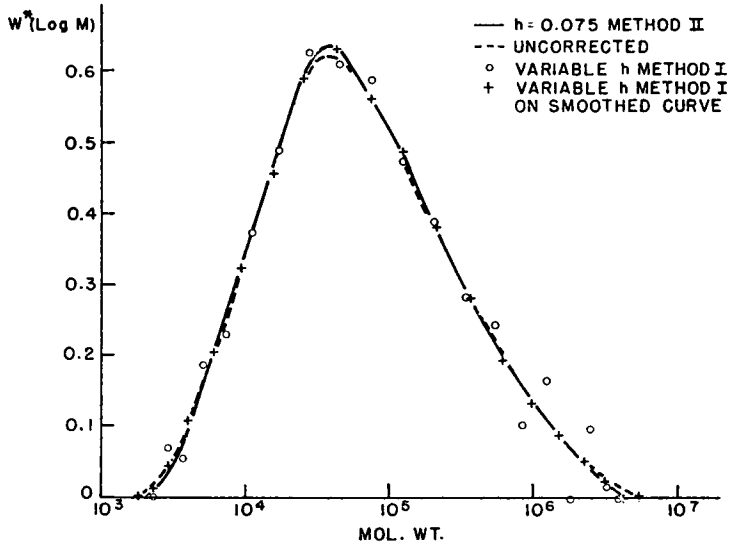


Fig. 2. Molecular weight distribution curves of example I.

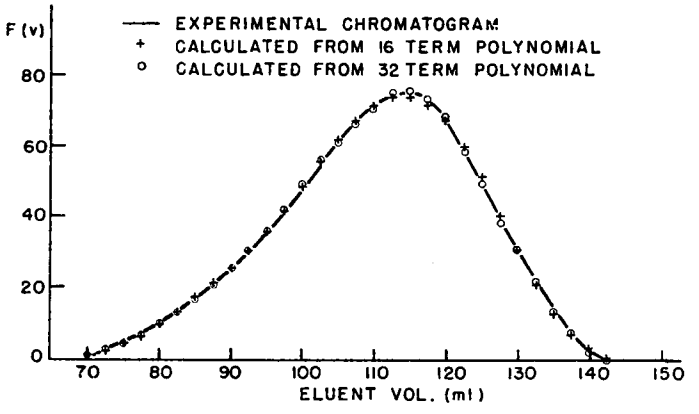


Fig. 3. Chromatogram of example I.

When $n = 16$ the oscillation disappeared but agreement between $F(v)$ calculated from the optimized $W(y_i)$ and the original $F(v)$ was unacceptable. At $n = 24$ the oscillation was even more severe than at $n = 20$.

Program II was tried next. The resolution factor $h$ used for the calculation was that corresponded to the peak of the chromatogram. The curve represented by the solid line in Figure 2 is the result. The fit between the polynomial calculated by Program II and the experimental chromatogram is shown in Figure 3.

The result of a third calculation is shown by the crosses in Figure 2. Program I was used in the third calculation using a variable $h$. The data used, however, were those smoothed by the 32-term polynomial shown in

TABLE I
$\bar{M}_w/\bar{M}_n$ Ratios for Example I

| Method of calculation | $\bar{M}_w/\bar{M}_n$ |
|---|---|
| Uncorrected | 7.56 |
| Program I, $n = 20$ | 7.16 |
| Program I, $n = 24$ | 7.14 |
| Program II, $n' = 32$ | 7.12 |
| Program I on smoothed data | 7.00 |

Figure 3. As the agreement of this polynomial with the chromatogram is well within the experimental errors, the crosses in Figure 2 should represent the true molecular weight distribution of the sample. The weight-average to number-average molecular weight ratios, $\bar{M}_w/\bar{M}_n$, obtained from these calculations are shown in Table I.

A few observations can be made from these calculations. (*1*) At both ends of a chromatogram where data cannot be read with high precision, the best solution calculated by program I is often not rational. When the data are smoothed by the polynomial of program II, program I then gives good results. (*2*) The crosses and the solid line curve in Figure 2 are almost indistinguishable. Thus the variation of $h$ with eluent volume can be ignored, a fact that justifies the use of Program II for treating broad-distribution chromatograms even when $h$ is not constant. (*3*) The difference shown in Figure 2 between the distribution corrected for resolution and that uncorrected is minor for the present broad-distribution sample. (*4*) It has been shown earlier[2] that if the chromatogram is representable by the polynomial in eq. (4), the average molecular weights can be calculated by the following equations:

$$\bar{M}_n = M_0 \Big/ \left\{ N e^{1/(4C_2^2p^2)} \sum_{i=0}^{n'} R_i \sum_{j=0}^{i} {}_iC_j Q'_{i-j} \left( \frac{1}{2C_2^2p^2} \right)^j \right\} \tag{6}$$

$$\bar{M}_w = M_0 N e^{1/(4C_2^2p^2)} \sum_{i=0}^{n'} R_i \sum_{j=0}^{i} {}_iC_j Q'_{i-j} \left( \frac{-1}{2C_2^2p^2} \right)^j \tag{7}$$

The symbols in these equations were defined in the earlier paper. For the two $n'$ values tried for this example, these equations were found to be useable only when $n' = 16$ but not when $n' = 32$. The average molecular weights for $n' = 32$ shown in Table I were calculated numerically with $v_a$ and $v_b$ as the limits of integration. This inconsistency indicates that the polynomial (32-term) for the solution $W(y)$ has nonzero values outside the limits of $v_a$ and $v_b$. In spite of this nature, $W(y)$ still represents the distribution correctly in the region of interest.

### Example II: Single Peak, Narrow Distribution

The resolution and molecular weight calibration for this example are the same as those for example I. Figure 4 shows the experimental chromato-

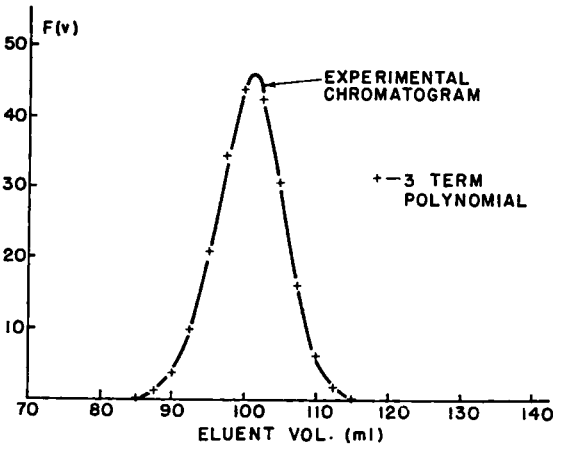


Fig. 4. Chromatogram of example II.

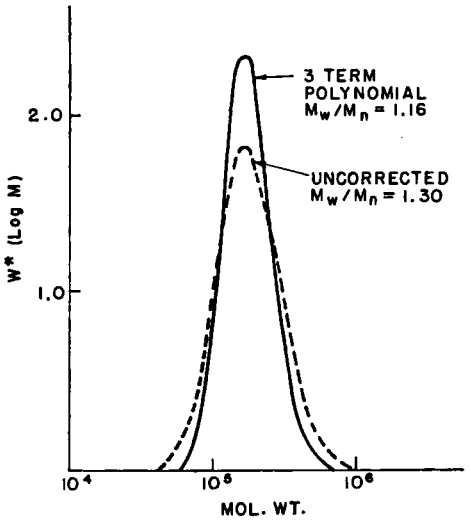


Fig. 5. Molecular weight distribution curves of example II.

gram and the polynomial calculated by program II with $n' = 3$. Figure 5 shows the uncorrected distribution and the distribution calculated by program II with $n' = 3$ and an $h$ which corresponded to the eluent volume at the peak. The correction for resolution shown in Figure 5 is much more significant than that for the previous example. Program I was tried on both the unsmoothed and the smoothed data. Both gave irrational results.

In this example, although a slightly better fit between the polynomial and the experimental chromatogram was obtained when $n'$ was greater than 3, the distribution calculated was irrational. In many cases of still narrower distribution the chromatogram can only be approximated by the

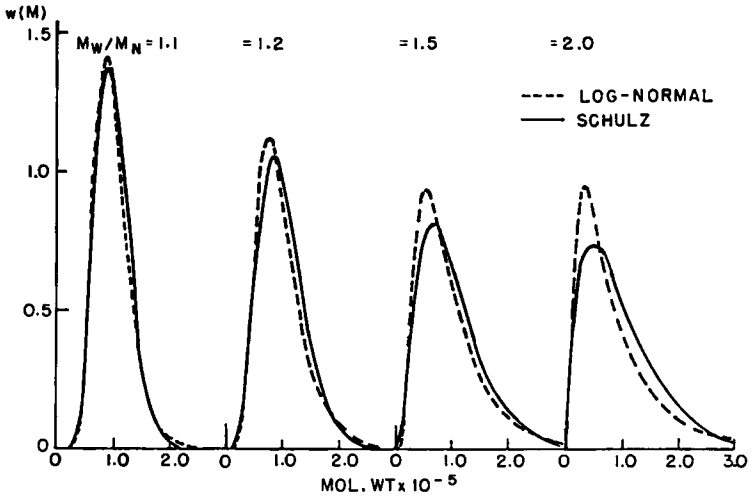


Fig. 6. Comparison of the log-normal and the Schulz distribution.

Gaussian function ($n' = 0$, 1, or 2). It appears then that the correction for resolution is more difficult to apply when the need is the most. However, for many narrow distribution chromatograms the Gaussian function is not a bad approximation. The reason is that many distribution functions become alike when the distribution is narrow. The comparison of log-normal distribution and the Schulz distribution[4] at different $\bar{M}_w/\bar{M}_n$ ratios shown in Figure 6 illustrates this point. The log-normal distribution, eq. (8), is the form for $W(y)$ when the chromatogram is Gaussian.

$$w(M) = (1/\beta\sqrt{\pi})(1/M) \exp \{-[\ln(M/M_0)]^2/\beta^2\} \quad (8)$$

The Schulz distribution, eq. (9), is known to correlate well the molecular weight distribution of many free radical-polymerized vinyl polymers and many condensation polymers.

$$w(M) = (-\ln \alpha)^b + {}^1M^b + {}^1\alpha^M/\Gamma(b + 2) \quad (9)$$

The symbols $\beta$, $M_0$, $b$, and $\alpha$ are adjustable parameters.

### Example III: Two Peaks, Broad Distribution

The resolution factor $h$ of this example and the three other examples to follow is shown by curve $B$ in Figure 1. The molecular weight–eluent volume relationship for these four examples is given by

$$v = 204.29 - 18.83(\log M) \quad (10)$$

Figure 7 shows the experimental chromatogram and the 31-term polynomial calculated by program II. The distribution calculated by program II by use of an $h$ which corresponded to the eluent volume midway between the two peaks is shown in Figure 8. Because of the poor fit at the peaks in Figure 7 the distribution calculated is distorted at the peaks.

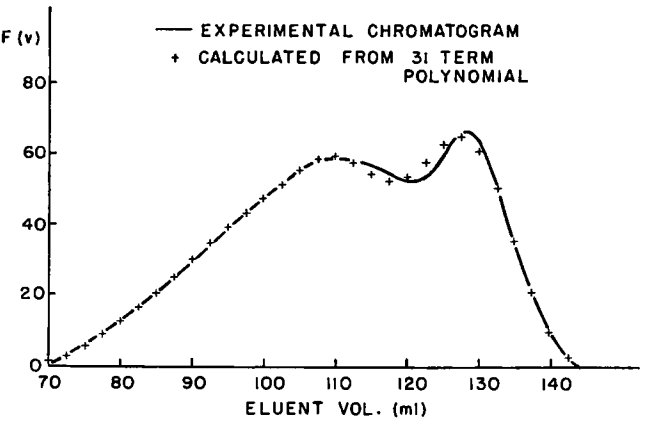


Fig. 7. Chromatogram of example III.

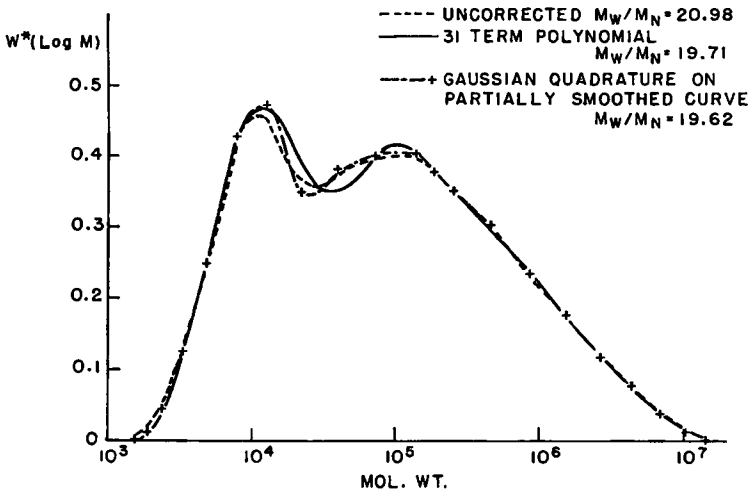


Fig. 8. Molecular weight distribution curves of example III.

An attempt was made to compute the distribution by program I using the original data. Like example I, the calculated result oscillated at both ends of the distribution curve. A final calculation was made using program I and the original data at the peaks but smoothed data at the ends of the chromatogram where good fit existed between the polynomial and the experimental chromatogram. This result is shown by the crosses in Figure 8 and they should represent points closest to the true molecular weight distribution of the sample.

As in the case of broad single-peak distribution, the correction of resolution for this example is minor. The agreement between the distributions calculated by program I and program II at the two ends shows again that the variation of $h$ with eluent volume can be ignored. The $\bar{M}_w/\bar{M}_n$ ratios calculated by the two programs agree also closely. Thus the distortion

at the peaks of the curve calculated by program II did not influence appreciably the $\bar{M}_n/\bar{M}_w$ ratio.

### Example IV: Two Peaks, Narrow Distribution

As shown in Figure 9 a good fit between the experimental chromatogram and a 31-term polynomial calculated by program II was obtained for this example. The distribution calculated by using an $h$ which corresponded to the eluent volume midway between the two peaks is shown in Figure 10. The correction for resolution for this case is again large as in the case of single-peak narrow distribution.

Attempts were made to calculate the distribution by program I with both the original data and the smoothed data. Neither results were satisfactory.

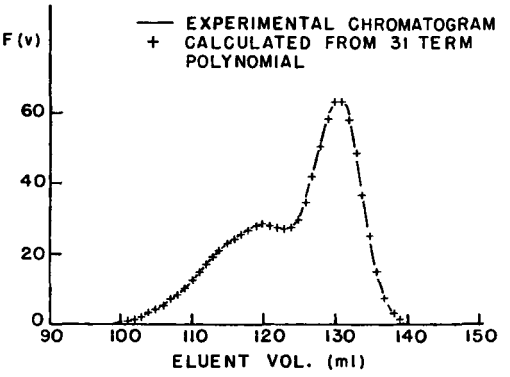


Fig. 9. Chromatogram of example IV.

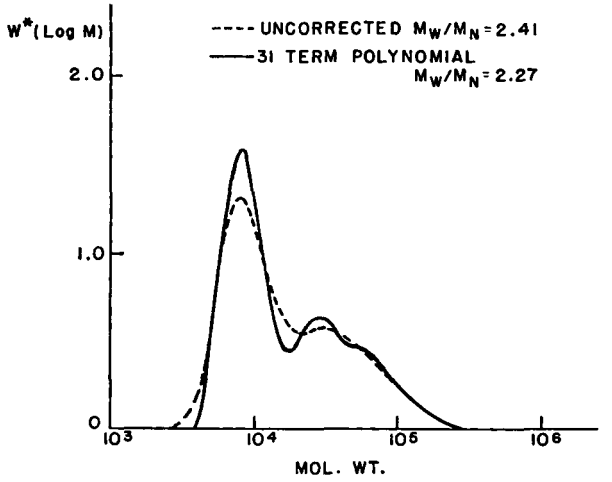


Fig. 10. Molecular weight distribution curves of example IV.

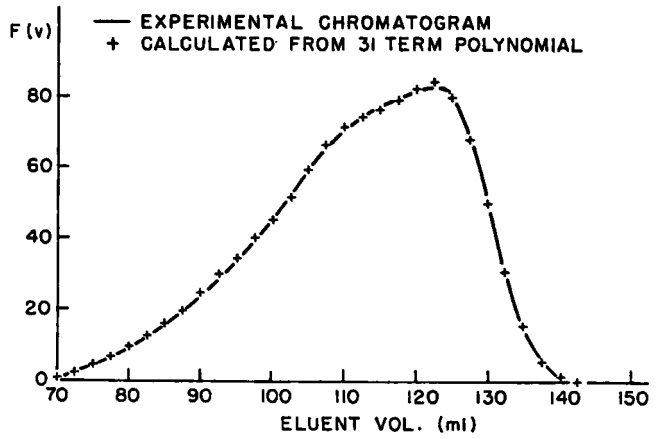


Fig. 11. Chromatogram of example V.

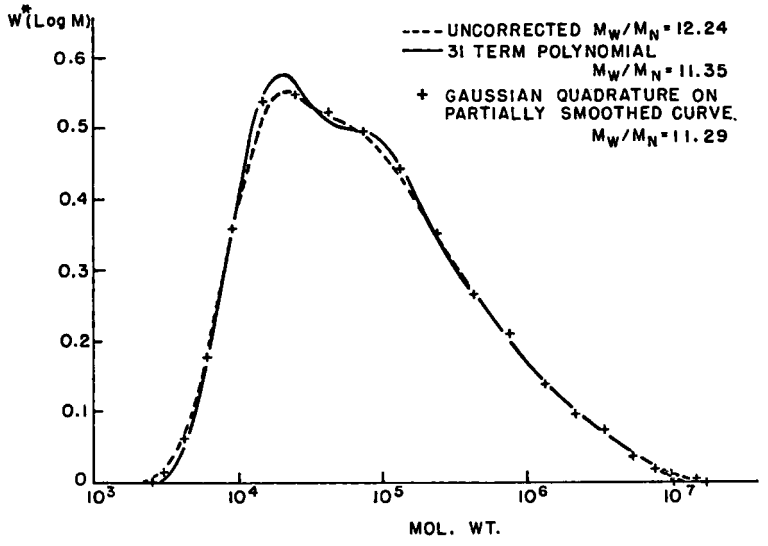


Fig. 12. Molecular weight distribution curves of example V.

## Example V: Broad Peak Distribution

Figure 11 shows the experimental chromatogram of this sample and the fit of a 31-term polynomial computed by program II. Figure 12 shows the distribution calculated from program II using an $h$ which corresponded to the eluent volume at the peak, the distribution calculated by program I for partially smoothed data points, and the uncorrected distribution.

The fit shown in Figure 11 appears to be reasonable. However, the slight deviation of the fit at the peak made the appearance of the two peaks more apparent in the distribution calculated by program II than that calculated by program I. The extent of the deviation shown in Figure 11 is probably

in the same order of the errors involved in the experimental reproducibility and the human ability in reading the chromatogram. This example therefore gives an indication of the uncertainty in determining the shape of a distribution.

### Example VI: Minor Peak at One End of a Broad Distribution

As shown in Figures 13 and 14 such a chromatogram cannot be fitted satisfactorily by the polynomial of program II nor can it be treated satisfactorily by program I. The exact molecular weight distribution of this sample therefore cannot be determined by the present calculation methods.

The $\bar{M}_w/\bar{M}_n$ ratio (shown in Fig. 14) calculated by program I is obviously too high; that calculated by program II is closer to the true value. In

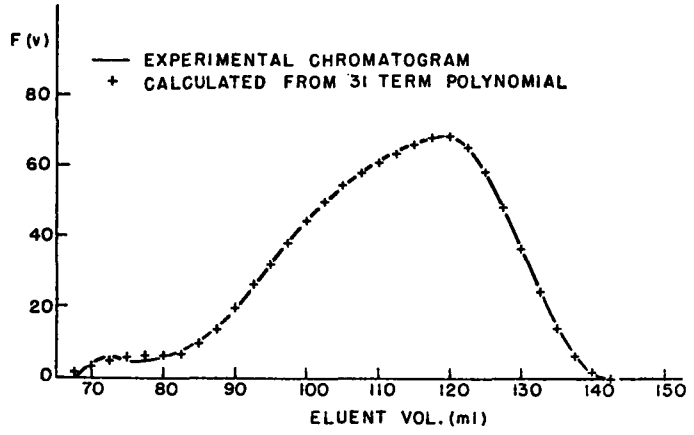


Fig. 13. Chromatogram of example VI.

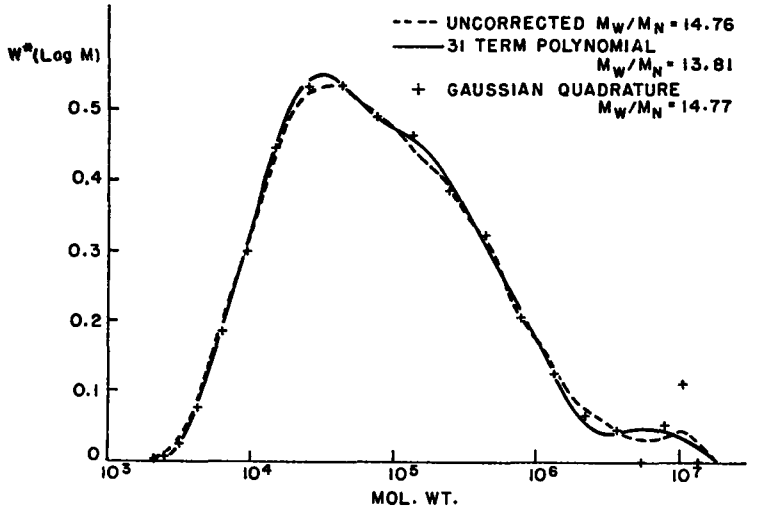


Fig. 14. Molecular weight distribution curves of example VI.

example I the $\bar{M}_w/\bar{M}_n$ ratio calculated by program I on unsmoothed data was correct, although the distribution oscillated at the ends.

## DISCUSSIONS

The above examples show that the importance of the correction of resolution depends on the distribution of the samples; less for broad distribution and more for narrow distribution.

In the application of the present calculation methods, whenever the experimental chromatograms can be fitted by the polynomial of eq. (4), program II is suitable even when $h$, the resolution factor, varies with the eluent volume. Generally, the larger the number of terms $n'$ used in the polynomial, the better the fit. For narrow distributions, however, the best fit may yield irrational molecular weight distributions and smaller $n'$ must be used. One may be limited to $n' = 1$ (or equivalently $n' = 0$, $n' = 2$) for some very narrow molecular weight distribution samples.

For complex chromatograms a good fit by the polynomial sometimes is not possible. Program I then must be used if one does not wish to ignore the correction for resolution. Program I in its present form often produces results that oscillate at the ends of the distribution. This oscillation can be removed in most cases by substituting the original data points at the ends of the chromatogram with points smoothed by the polynomial if there is good fit between the polynomial and the chromatogram in these regions. Like example VI there are occasional cases the chromatograms are so complex that none of the two programs can be used satisfactorily.

A reason for the above difficulties is that eq. (1) requires a certain consistency in the input data for $W(y)$ to be positive and nonoscillatory. This consistency of data often demands more significant digits than can be read from the experimental chromatograms. In program II the data points are smoothed by the polynomial before the actually calculation and this demand becomes less critical. In program I the experimental data are used directly and consequently it is more troublesome to use. Another source of the data inconsistency is the distortion of the chromatogram from the ideal shape due to experimental errors. The narrower the distribution, the more stringent is this requirement.

It is unlikely that the above condition is caused by the assumption of the Gaussian function to represent the broadening effect of imperfect resolution. Rather, it is difficult to avoid this small deviation from perfect consistency in the experimental chromatogram no matter how perfect the assumed function describes the ideal broadening effect. It should be noted that program I can be easily modified to accommodate other functions for this broadening effect beside the Gaussian function used in the present work. Program I can also be modified so that it is less sensitive to the inconsistency of input data. This modification has not been made because most of the chromatograms encountered by us can be adequately treated by the existing programs, and their accuracy is consistent with the experimental errors of the GPC units.

The author is indebted to two of his colleagues at The Dow Chemical Company for their assistance in this work. M. Klein of the Computations Research Laboratory wrote the linear programming part of program I and contributed many valuable suggestions in the use of this program. G. W. Knight of Pilot Plant Department, Texas Division, furnished all the chromatograms and the calibration data for the present examples.

## References

1. M. Hess and R. F. Kratz, *J. Polymer Sci.*, in press.
2. L. H. Tung, *J. Appl. Polymer Sci.*, **10**, 375 (1966).
3. L. H. Tung, J. C. Moore, and G. W. Knight, *J. Appl. Polymer Sci.*, **10**, 1261 (1966).
4. G. V. Schulz, *Z. Physik. Chem.*, **B43**, 25 (1939).

## Résumé

Six échantillons ont été utilisés pour illustrer l'application de la méthode rapportée antérieurement concernant le calcul de la distribution des poids moléculaires au départ de chromatogrammes par perméation sur gel. Ces exemples montrent que la correction pour une résolution imparfaite de la colonne GPC est importante lorsque la distribution est étroite mais moindre lorsque la distribution est large. Ces exemples montrent également que la variation du facteur de résolution $h$ (défini précédemment) avec le volume éluant peut être négligé dans ce calcul. Lorsque les chromatogrammes sont très étroits de distribution ou lorsqu'ils sont complexes, on rencontre certaines difficultés. Les deux programmes de calcul en vue d'effectuer les calculs numériques, décrits précédemment, sont adéquats dans le cas de ces difficultés mais sont également soumis à certaines limitations.

## Zusammenfassung

Eine früher angegebene Methode zur Berechnung der Molekulargewichtsverteilung von Gelpermeationschromatogrammen wurde an sechs Beispielen erläutert. Diese Beispiele zeigen, dass die Korrektur für die unvollständige Auflösung der GPC-Säule bei enger Verteilung wichtig ist, bei breiter Verteilung jedoch geringe Bedeutung hat. Sie zeigen auch, dass die Abhängigkeit des früher definierten Auslösungsfaktors $h$ vom Elutionsvolumen bei der Berechnung vernachlässigt werden kann. Bei Chromatogrammen mit sehr enger Verteilung oder bei komplexen Chromatogrammen treten gewisse Schwierigkeiten auf. Zwei zur Ergänzung der früher beschriebenen numerischen Berechnungen geschriebene Computerprogramme erweisen sich für diese schwierigen Falle als adäquat, aber auch hier treton Begrenzungen auf.

Received April 14, 1966
Prod. No. 1382