JOURNAL OF APPLIED POLYMER SCIENCE VOL. 18, PP. 1549–1556 (1974)

Molecular Weights by Computer from Gel Permeation Data on Punched Tape

CHARLES L. SWANSON, JOSEPH O. ERNST, and LEWIS A. GUGLIEMELLI, *Northern Regional Research Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Peoria, Illinois 61604*

Synopsis

A Fortran IV program has been written that reduces digitized data from gel permeation chromatography to number and weight average molecular weights. This computer program is suitable for routine calculations whenever corrections are unnecessary for axial diffusion and branching effects.

INTRODUCTION

Manual computation of molecular size and weight data from raw molecular weight distribution (MWD) curves obtained by gel permeation chromatography (GPC) is tedious and subject to error. Accuracy in such calculations is essential when interpreting effects of experimental treatments on polymer molecular weight or structure from GPC data.

Computer processing of GPC data automatically recorded on teletype punched tape greatly increases the accuracy of molecular weight calculations because many more data points can be used than are practical manually and because errors in determining baselines and curve heights are reduced. Most programs for processing taped GPC data are designed to

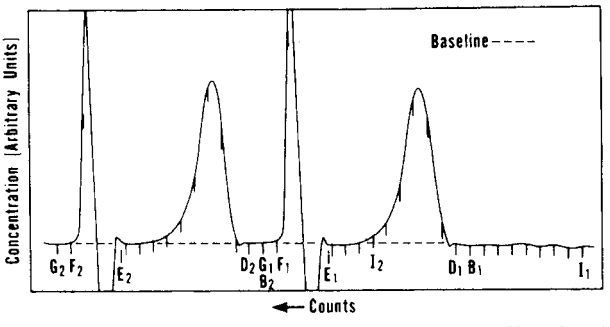

Fig. 1. Gel permeation chromatography (GPC) elution profiles showing curve segments ($\overline{B_iD_i}$ and $\overline{F_iG_i}$) selected for baseline determination and a segment ($\overline{D_iE_i}$) for calculation of molecular weight data. I_i indicates sample injection, and each count is indicated by a small vertical line under the curve.

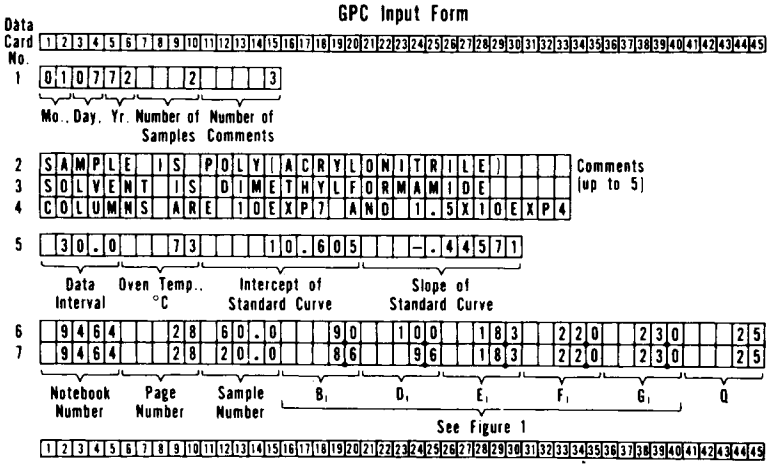

Fig. 2. Raw data input form showing statements that control sorting of the taped data. Data cards are keypunched in these formats.

plot MWD after correcting for axial spreading, other chromatographic phenomena, or long- and short-chain branching effects.[1–5] Such complex procedures are often unwarranted in routine polymer characterizations. Other programs compute number-average molecular weight ($\bar{M}_n$), weight-average molecular weight ($\bar{M}_w$), viscosity-average molecular weight, z-average molecular weight, viscosity, and differential and integral MWD's, but are designed for data input from card decks rather than teletype tape.[6] Our program computes $\bar{M}_n$, $\bar{M}_w$, and d values directly from taped GPC curve data in the output form provided by the digital translator of Waters Associates, Inc. Computations follow the example of Harmon.[7]

PROGRAM DESIGN

Data on each tape include the gross curve heights measured at uniform data intervals (DI's) and the times of each volume count (C) or sample injection event (percentage of DI elapsed since preceding height measurement) along with their respective identifying codes. A strip chart record of typical data is shown in Figure 1. Note that the baseline drifts and must be redetermined for each new curve; solvent impurity data follow each sample's data and must be excluded from calculations; and the succeeding sample may be injected before the preceding sample is completely eluted, provided the elution curve can return to the baseline between peaks.

Specific details required for sorting and computing the data on each tape are supplied to the computer by an input subdeck (Fig. 2) that lists both the number of samples on the tape (NS) and their sample identifying codes; the count intervals where baseline and peak data are to be calculated ($\overline{B_iD_i}$ and $\overline{F_iG_i}$ for baseline and $\overline{D_iE_i}$ for peak, Fig. 1); the DI; the weight of a unit length of polymer (Q); and the slope and intercept fo the equation that relates molecular contour length (A) to C. The general subscript i assumes the value 1 for the first sample injected, 2 for the second sample,

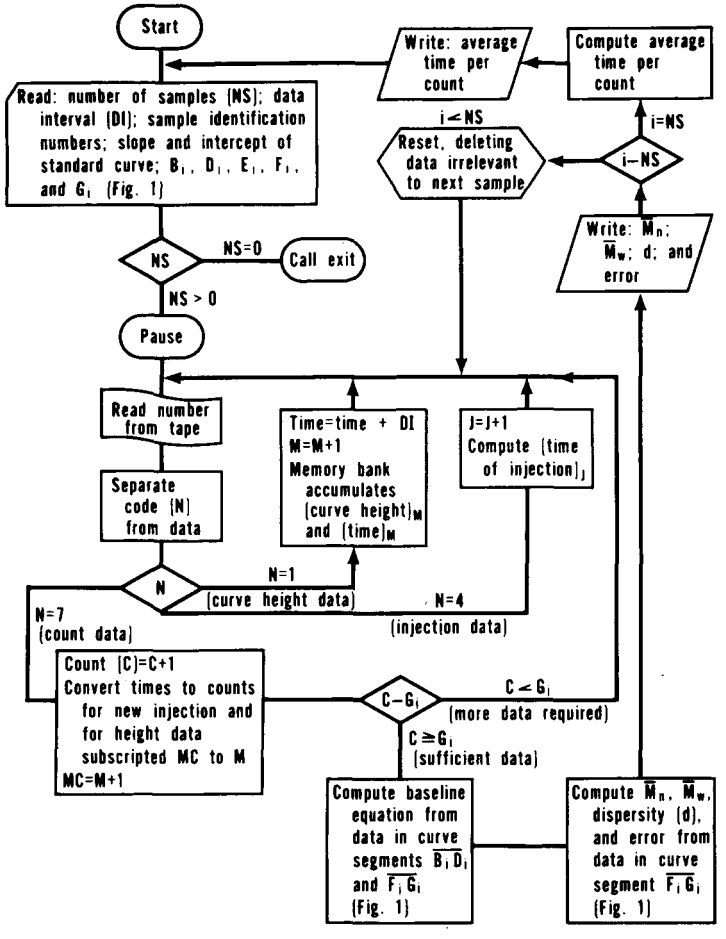

Fig. 3. GPC program flow chart. Subscript i corresponds to the position of a sample in the sequence of injections (e.g., i equals 3 for the third sample).

etc. Data on the tape are sorted as outlined in the abbreviated flow chart (Fig. 3) until count G_1 is located. Then the equation of the baseline is computed by the method of least squares. From this baseline equation, the equation of the standard curve ($\log A_{\text{std}} = mC + b$) and the height and count data in the computer memory for curve segment $\overline{D_1E_1}$, the computer determines the sum of the net curve height (ΣH) and the sums $\Sigma(H/A)$ and $\Sigma(H \cdot A)$ for the sample. From these sums it computes

$$A_n = \Sigma H/\Sigma(H/A)$$

$$A_w = \Sigma(H \cdot A)/\Sigma H$$

$$\bar{M}_n = Q \cdot A_n$$

$$\bar{M}_w = Q \cdot A_w$$

and

$$d = A_w/A_n.$$

After printing the molecular weight data for the sample, the computer discards information that is irrelevant to the succeeding sample and continues. When calculations for the first tape are completed, the computer reads another card subdeck and pauses until a new tape is loaded onto the reader. The program continues until the computer finds $NS = 0$ (a blank #1 card).

```
// JOB

LOG DRIVE   CART SPEC   CART AVAIL  PHY DRIVE
  0000        1816        1816        0000

V2 M06   ACTUAL  8K  CONFIG  8K

// FOR
* LIST SOURCE PROGRAM
* ONE WORD INTEGERS
* IOCS ( CARD , 1132 PRINTER , TYPEWRITER )
      DIMENSION M1(350),CTM1(350),CNTI(3),NBN(20),NPN(20),SPL(20),CBL1(2
     10),CBL2(20),CSP(20),CBL3(20),CBL4(20),JIC(3),IALPH(30,5),Q(20)
      M1MAX=350
      JIMAX=3
  100 KL=0
      M1=0
      M2=0
      IX=1
      IDP=0
      TIMEC=0
      TIMED=0.
      TPC=0
      JI=0
      JC=0
      K=1
      KLM=1
      JK=1
      TTPC=0.
      READ (2,120) MO,MDAY,MYR,NS,NC
  120 FORMAT (3I2,I4,I5)
      IF(NS)130,130,140
  130 CALL EXIT
  140 WRITE (3,160) MO,MDAY,MYR
  160 FORMAT (' MOLECULAR SIZE BY GEL PERMEATION CHROMATOGRAPHY'4XI2'/'I
     12'/'I2/'0')
      READ (2,200) ((IALPH(I,J),I=1,30),J=1,NC)
  200 FORMAT (30A2)
      WRITE (3,220) ((IALPH(I,J),I=1,30),J=1,NC)
  220 FORMAT (1X,30A2)
      READ (2,260) DI, OVT,YSTD,SLSTD
  260  FORMAT (2F5.0,2F10.0)
      WRITE (3,270) YSTD,SLSTD
  270 FORMAT (' STANDARD CURVE IS   LOG(A)='F10.4,F10.4'(C)')
      WRITE (3,280) OVT
  280 FORMAT (' OVEN TEMPERATURE IS 'F5.1' DEGREES CENTIGRADE')
      WRITE (3,320)
  320 FORMAT ('0      SAMPLE          M-SUB-N    M-SUB-W        D        ERR
     1OR INTERCEPT      SLOPE     Q')
      READ (2,360) (NBN(I),NPN(I),SPL(I),CBL1(I),CBL2(I),CSP(I),CBL3(I),
     1CBL4(I),Q(I),I=1,NS)
  360 FORMAT (2I5,6F5.1,F6.1)
      PAUSE 3001
      LIN=-1
  400 CALL TLTAP (NO,NUMBR)
      IF(NO-4)420,500,460
  420 IF(NO-3)460,440,460
  460 WRITE (1,480) LIN,NO,NUMBR
      PAUSE 3003
      GO TO 400
  440 LIN=NUMBR
```

(a)

Fig. 4 (*continued*)

PROGRAM OPERATION

An input subdeck (Fig. 2) must be prepared for each tape. Each input subdeck includes (in the order listed): A card giving the date, the number of samples on the tape, and the number of informational comments (up to five) to be read and printed in the output; the comment cards (each contain-

```
      GO TO 400
480   FORMAT (' ERROR LINE='I6'  NO='I2'  VALUE='I6)
500   N=NUMBR/1000
      NUM=NUMBR-1000*N
      IF(N-1)460,600,520
520   IF(N-4)460,690,540
540   IF(N-7)460,740,560
560   IF(N-9)460,980,460
600   IDP=IDP+1
      TIME=IDP*DI
      IF(JI-1)400,660,660
660   M1=M1+1
      IF(M1-M1MAX)680,680,460
680   M1(M1)=NUM/10.
      CTM1(M1)=TIME
      GO TO 400
690   IF(JI)695,695,700
695   IF(TPC)696,696,701
696   WRITE (3,697)
697   FORMAT (' SAMPLE WAS INJECTED WITHOUT TWO PRIOR COUNTS')
      JK=2
      GO TO 400
700   IF(JIC(JI)-JC)701,400,400
701   JI=JI+1
      IF(JI-JIMAX)720,720,460
720   CNTI(JI)=TIME+NUM*DI/100.
      JIC(JI)=JC
      COAJ=1-(CNTI(JI)-TIMEC)/TPC
      M2=M1
      GO TO 400
740   JC=JC+1
      KI=M1
      TIMEC=TIME+NUM*DI/100.
      IF(TIMED)815,815,760
760   TPC=TIMEC-TIMED
      TTPC=TTPC+TPC
      IF(KI)815,815,764
815   TIMED=TIMEC
      IX=M1+1
      GO TO 400
764   IF(KL-1)782,765,765
782   DO 783 MX=IX,KI
783   CTM1(MX)=(CTM1(MX)-CNTI(1))/TPC
      CAJ=(TIMEC-CNTI(1))/TPC
      CNTI(1)=0
      KL=1
      GO TO 815
765   DO 780 MX=IX,KI
780   CTM1(MX)=(CTM1(MX)-TIMED)/TPC+JC-2.-JIC(1)+CAJ
      KMI=JIC(JI)+1
      IF(KMI-JC)820,810,820
810   CNTI(JI)=(CNTI(JI)-TIMED)/TPC+JC-2.-JIC(1)+CAJ
820   TIMED=TIMEC
      IX=M1+1
      CNT=CTM1(M1)
      CB4=CBL4(JK)
      IF (CNT-CB4) 400,830,830
830   MIX=M1
      CB1=CBL1(JK)
      CB2=CBL2(JK)
      CS=CSP(JK)
      CB3=CBL3(JK)
```

(b)

Fig. 4 (*continued*)

ing up to 60 characters and spaces); a card listing the data interval, oven temperature, and the intercept and the slope of the standard equation; a card for each sample, which lists its identification (notebook number, page number, and sample number), B_i, D_i, E_i, F_i, and G_i (from the GPC

```
      DO 877 M=1,MIX
      COUNT=CTH1(M)
      IF (COUNT-CB1) 840,850,850
  840 MLOW=M+1
      GO TO 877
  850 IF (COUNT-CB2) 860,860,870
  860 MMID=M
      GO TO 877
  870 IF (COUNT-CS) 875,880,880
  875 MHIGH=M
  877 CONTINUE
  880 DO 890 M=1,MIX
      COUNT=CTH1(M)
      IF (COUNT-CB3) 881,881,887
  881 MLB1=M+1
      GO TO 890
  887 IF (COUNT-CB4) 888,888,900
  888 MLB2=M
  890 CONTINUE
  900 SCNT=0.
      SH=0.
      SCNTH=0.
      SCNT2=0.
      DO 910 M=MLOW,MMID
      SCNT=SCNT+CTH1(M)
      SH=SH+H1(M)
      SCNTH=SCNTH+CTH1(M)*H1(M)
  910 SCNT2=SCNT2+CTH1(M)**2
      DO 915 M=MLB1,MLB2
      SCNT=SCNT+CTH1(M)
      SH=SH+H1(M)
      SCNTH=SCNTH+CTH1(M)*H1(M)
  915 SCNT2=SCNT2+CTH1(M)**2
      Z=MMID+MLB2+2-MLB1-MLOW
      SLOPE=(SCNT*SH-Z*SCNTH)/(SCNT*SCNT-Z*SCNT2)
      YCEPT=(SH-SLOPE*SCNT)/Z
      SHXA=0.
      SHDA=0.
      SHCRV=0.
      DO 940 M=MMID,MHIGH
      HCRV=H1(M)-SLOPE*CTH1(M)-YCEPT
      IF(HCRV)920,930,930
  920 HCRV=0.
  930 ASTD=SLSTD*CTH1(M)+YSTD
      ASTD=EXP(ASTD*2.30258509)
      SHXA=SHXA+ASTD*HCRV
      SHDA=SHDA+HCRV/ASTD
  940 SHCRV=SHCRV+HCRV
      ABARN=SHCRV/SHDA
      ABARW=SHXA/SHCRV
      HBARN=Q(JK)*ABARN
      HBARW=Q(JK)*ABARW
      DSPTY=ABARW/ABARN
      ERROR=H1(MHIGH)-SLOPE*CTH1(MHIGH)-YCEPT
      WRITE (3,950) NBN(JK),NPN(JK),SPL(JK),HBARN,HBARW,DSPTY,ERROR,YCEP
     1T,SLOPE,Q(JK)
  950 FORMAT (1X I5,' -',I3,' -'F6.1,2F10.0,2F10.2,F10.4,F10.3,F6.1)
      IF (JK-NS) 960,980,980
  960 IF (M2) 964,964,968
  964 M1=0
      J1=0
      KL=0
```

(c)

Fig. 4 (*continued*)

```
      IX=1
      JK=JK+1
      GO TO 400
  968 L2=M1-M2
      L3=M2+1
      DO 970 IM=1,L2
      CTM1(IM)=CTM1(L3)-CNTI(2)
      M1(IM)=M1(L3)
  970 L3=L3+1
      IX=IX-M2
      M1=M1-M2
      M2=0
      CNTI(1)=CNTI(2)
      JIC(1)=JIC(2)
      JI=1
      KL=2
      CAJ=COAJ
      JK=JK+1
      GO TO 400
  980 AVTPC=TTPC/(JC-1)
      WRITE (3,990) AVTPC
  990 FORMAT (// '0AVERAGE TIME PER COUNT IS 'F6.2,' SEC')
      GO TO 100
      END

FEATURES SUPPORTED
 ONE WORD INTEGERS
 IOCS

CORE REQUIREMENTS FOR
 COMMON       0  VARIABLES   2002  PROGRAM   1696

 END OF COMPILATION

// XEQ
```

(d)

Fig. 4. GPC program listing.

strip chart, Fig. 1), and the Q factor to be used in the calculations. All numbers must be right-justified integers unless shown in floating point (decimal) form.

For computation, the input subdecks are loaded into an IBM card reader on top of the program deck (listed in Fig. 4) in the order that the tapes will be processed, and two blank cards are put on the back of the entire deck. The first tape is loaded onto an IBM 1134 paper tape reader. Each tape must be loaded so that the computer finds two counts before it locates the first injection since the time of the first injection is calculated from the time between the preceding two counts. (If less than two counts are observed before injection, the computer will indicate this fact on the printer and proceed to the next sample.)

After the program is loaded, the computer reads the first subdeck, prints the date and comments (Fig. 5), and pauses until the "program start" button is pushed. As $\bar{M}_n$, $\bar{M}_w$, and d are calculated for each sample, they are listed along with error, baseline, intercept, slope, and Q factor according to the format in Figure 5. Error is the difference in heights of the baseline and sample curve at count E_t. Deviation of this residual from zero by more than 0.5 unit for a sample implies nonvalidity of its molecular weight data. Intercept and slope define the equation of the baseline for that sample. After all samples on a tape have been processed, the average time

```
MOLECULAR SIZE BY GEL PERMEATION CHROMATOGRAPHY    10/20/72

SAMPLE IS POLY(ACRYLONITRILE)
SOLVENT IS DIMETHYLFORMAMIDE
DATA INTERVAL IS 30 SEC.
COLUMNS ARE 10EXP7 AND 1.5X10EXP6
STANDARD CURVE IS   LOG(A)=   10.6050   -0.4457(C)
OVEN TEMPERATURE IS  73.0 DEGREES CENTIGRADE

     SAMPLE          M-SUB-N   M-SUB-W       D      ERROR INTERCEPT     SLOPE    Q
 9535 -  5 -  1.1  2479055.  8030165.     3.23    -0.81    15.7993     0.066  25.0
 9535 -  6 -  1.0  2162225.  7020786.     3.24    -0.18    16.6738     0.044  25.0

AVERAGE TIME PER COUNT IS 289.68 SEC
```

Fig. 5. Computer output from GPC program.

elapsed per count is computed and printed, and the card reader processes another input subdeck. The computer then pauses until a new tape is loaded onto the tape reader and the program start button is depressed again.

DISCUSSION

The program can be modified readily to process information from three GPC units recorded on a single tape. Only slight modification is required to use a nonlinear relationship between A or M and C as the standard equation. Calibration of GPC with samples of the test polymer that have small d values and known peak M's would remove the uncertainties associated with use of the Q factor.

TLTAP (statement 400, Fig. 4) refers to a tape reading subroutine. Specific inquiries about this subroutine can be directed to the Northern Regional Research Laboratory in care of J. O. Ernst.

The mention of firm names or trade products does not imply that they are endorsed or recommended by the U.S. Department of Agriculture over other firms or similar products not mentioned.

References

1. K. S. Chang and R. Y. M. Huang, *J. Appl. Polym. Sci.*, **13**, 1459 (1969).
2. E. M. Rosen and T. Provder, *J. Appl. Polym. Sci.*, **15**, 1687 (1971).
3. M. Hess and R. F. Kratz, *J. Polym. Sci. A-2*, **4**, 731 (1966).
4. J. H. Duerksen and A. E. Hamielec, *J. Appl. Polym. Sci.*, **12**, 2225 (1968).
5. E. E. Drott and R. A. Mendelson, *J. Polym. Sci. A-2*, **8**, 1373 (1970).
6. H. E. Pickett, M. J. Cantow, and J. F. Johnson, *J. Appl. Polym. Sci.*, **10**, 917 (1966).
7. D. J. Harmon, *J. Polym. Sci. C*, **8**, 243 (1965).

Received November 2, 1973
Revised November 13, 1973