Thermochimica Acta, 14 (1976) 13–24

AUTOMATED AND COMPUTERIZED SYSTEM FOR PURITY DETERMINATION BY DIFFERENTIAL SCANNING CALORIMETRY*

S. A. MOROS AND D. STEWART

Quality Control Dept., Hoffmann-La Roche Inc., Nutley, N.J. 07110 (U.S.A.)

ABSTRACT

In order to evaluate and implement a widely used DSC method for purity determination, a commercial thermal analysis instrument was "interfaced" by means of an analog-to-digital converter and a communication terminal to a time-sharing computer. A computer program specifically designed for this system evaluates data for calibration purposes or calculates the purity. The program includes steps to detect the peak opening and closing, determine the area, evaluate the partial energies at various temperatures, linearize the $1/F$ vs. T plot, and calculate mole % purity. The analysis (or calibration) proceeds without intervention, or at the operator's control, provides complete versatility by accommodating a choice of instrumental or computational factors, including thermal lag, heat capacity, iteration increment, and the portion of the curve ($1/F$ limits) to be considered.

This automated and computerized system permits data accumulation and rapid purity calculation without human error, facilitating the systematic evaluation of theoretical and practical factors in the determination of purity.

INTRODUCTION

The objective of this work was to automate the thermal data acquisition and computerize the calculations employed in the determination of purity by differential scanning calorimetry[1,2], thereby eliminating sources of human error and reducing the time requirements. The computer program was designed to accommodate a wide variety of instrumental and computational factors.

SYSTEM DESCRIPTION

The system consists of three subsystems: (A) instrumentation for quantitatively sensing the thermal effect; (B) instrumentation for transmitting the data in a format compatible with: (C) the data processing subsystem.

---

*Presented at the 5th North American Thermal Society Meeting, Peterborough, Ontario, Canada, June 8–14, 1975.

(A) *Thermal sensing subsystem*

The thermal effect is detected by a Perkin-Elmer DSC 1b or II calorimeter. By means of this instrument experimental conditions such as sensitivity, scan rate, and starting temperature are selected and controlled. The energy requirements of a sample in the calorimeter are detected as it is heated up to and through its melting temperature at a particular rate. The energy is quantitatively expressed as an analog output signal.

(B) *Data acquisition subsystem*

Data acquisition involves the combined operation of four separate components which together function as an analog-instrument-to-digital-computer interface. These are an amplifier (Princeton Applied Research programmable differential amplifier, Model 265), an analog-to-digital converter (Princeton Applied Research analog-to-digital converter, Model 260), a teleprinter interface (Princeton Applied Research teleprinter/system interface, Model 262), and a teleprinter (Data Acquisition Systems teleprinter, Model 33, or Texas Instruments, Inc. Silent 700 ASR electronic data terminal).

(1) *Differential amplifier.* The amplifier functions to bring the 10 mV analog input signal to the optimum level for subsequent processing in an analog-to-digital converter. The amplifier is characterized by precise (±0.05%) automatic or manual gains of ×1, ×10 or ×100 with high (typically ±5 μV) input stability and common mode rejection over a dynamic range of 10^6:1 with inputs between ±14 V to any of nine channels.

(2) *Analog-to-digital converter.* The converter encodes analog signals with a full scale value of ±14 V into a digital "word" indicating the magnitude of the signal; the "word" consists of a polarity sign, an over-range digit, four data digits, and an exponential digit to convert to millivolts. Thus, an input of 2.345 V is converted to +0.2345E4 and transmitted in Binary Coded Decimal (BCD) form. Using the dual slope integrating technique for the conversion process, the converter is capable of full scale accuracy of +0.01% with a 200 msec conversion period. In addition to the BCD output, a front-panel NIXIE readout display provides a visual indication of the sign, overrange and data digits. The analog-to-digital conversion is initiated by means of an internal clock with selectable periods of from 20 msec to 20 sec to control the number of data points examined in a given experiment.

(3) *Teleprinter/system interface.* The teleprinter/system interface accepts the digital (BCD) data from the converter, formats it, encodes it into ASCII (American Standard Code for Information Interchange), and transmits the data to a teleprinter or a "modem" for retransmission to a remote computer. Controllable parameters include data transmission rate, the number of data words per line, and both the data-word (channel identification) and end-of-line formats.

(4) *Teleprinter.* The function of the teleprinter in this subsystem is two-fold. One is to provide a punched paper or magnetic tape of the digital form of the data in a given experiment and, if desired, a printed copy. The other function of the teleprinter

is to serve as a communication terminal with a time-sharing computer via telephone line using a data modem.

(C) *Data processing subsystem*

The final subsystem of this automated system is the computer and program for processing the accumulated data. Subsequent sections go into detailed consideration of the physical-chemical basis and the mathematical operations required for the calculation of purity. Once developed, the requirements were transmitted to a computer programming group where the program was written in FORTRAN for use with a time-sharing Honeywell H-6080 computer which performs the thermal purity calculation once the data have been placed in an intermediate disc-storage file and after the experimental conditions have been entered in a question-and-answer mode. The "debugging" of the program for errors arising from ambiguities in pertinent mathematical steps as well as from computer-related technical difficulties was accomplished by using synthetic data as described under Tests of System Operation.

DESCRIPTION OF THE DATA PROCESSING SUBSYSTEM

Basis of purity calculation

The purity determination is based on the Van't Hoff equation which relates the melting point depression to the concentration of impurity in the sample. The equation may be expressed as:

$$T_s = T_0 - \frac{RT_0^2 \chi_2}{\Delta H_f} \frac{1}{F}$$

where T_s = the instantaneous temperature of the sample;

T_0 = the melting point of the pure major component;

R = the gas constant;

χ_2 = the mole fraction of impurity;

ΔH_f = the heat of fusion of the major component;

F = the fraction of sample melted at T_s.

The equation has the form of a straight line; a plot of $1/F$ versus T_s should give a straight line with intercept T_0, and slope $-RT_0^2\, \chi_2/\Delta H_f$. Since ΔH_f can be determined from the total heat energy required to melt the sample and F is the ratio of the amount of heat energy required to melt part of the sample to the energy required to melt the entire sample, enough information is available from the energy–temperature curve to calculate the mole fraction of impurity, χ_2.

As this single equation does not account for the experimental observation made with a dynamic system, several "corrections" of the data are made at various stages (detailed below). As a result, the calculations become quite involved and the application of the method becomes impractical without a computer. Table 1 lists and defines symbols used in the following description of the individual operations required and Fig. 1 illustrates the portions of the thermal curve under discussion.

TABLE 1

DEFINITIONS AND DIMENSIONS OF THE PARAMETERS INVOLVED IN THE CALCULATIONS

Symbol	*Unit*	*Definition*
G	mcal sec^{-1} mV^{-1}	Range value of the instrument
w	mg	Weight of sample or standard
S	°C min^{-1}	Scan rate of the instrument
M	g mol^{-1}	Molecular weight
R	cal °C^{-1} mol^{-1}	Gas constant (1.9872)
ΔH_f	cal mol^{-1}	Heat of fusion
X_2	dimensionless	Mole fraction impurity
V_i	mV	DSC output value of data point i
F	dimensionless	Mole fraction melted at T_s
DR	sec	Time between data points
T_s	°C	Instantaneous temperature of sample
T_p	°C	Instantaneous program temperature
T_g	°C	Temperature at which data accumulation begins
T_0	°C	Melting temperature of pure sample
T_m	°C	Melting temperature of standard material
t	sec	Time from beginning of data accumulation
N_i	dimensionless	Number of data points accumulated at point i
L	mV sec^{-1}	Thermal lag of system
A_t	mV sec	Total area of thermal curve
A_p	mV sec	Partial area of thermal curve
a_c	mV sec	Correlation of partial area
C	mcal °C^{-1}	Heat capacity correction
Q	°C	Slope of best fit line in linearization process
k	dimensionless	Iteration correction

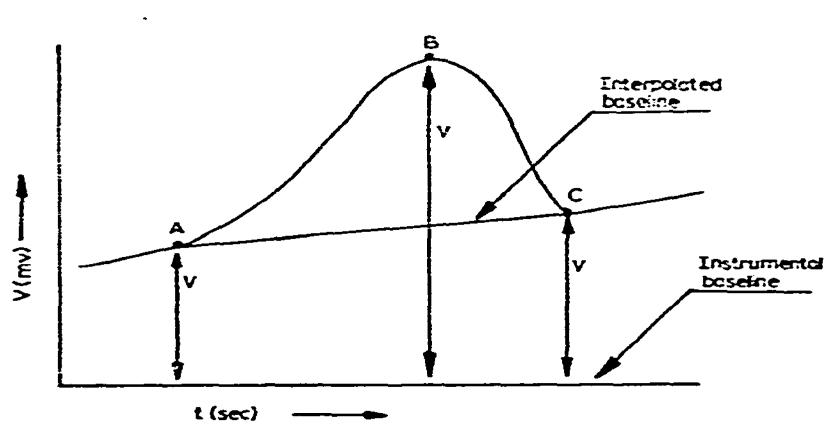

Fig. 1.

Peak sensing

Peak sensing requires that two points in the data list be detected: one that begins the peak (A) and one that ends the peak (C). This is accomplished for the peak opening by recording a substantial portion of the flat DSC curve preceding the peak, fitting the first 60 data points to a straight line, and examining each subsequent data point for conformance. Once a point has been found to lie outside a 2-sigma control level a check of each of the next 30 data points is made; if deviations outside the control level continue, a peak opening has been sensed. The last point within the 2-sigma criterion is the peak start (A). Sensing for the end of the peak is the same as for the beginning, except that the data are examined in reverse order.

Baseline construction

To obtain the total area under the thermal curve (an area bounded by the peak curve itself and a line connecting "peak start" and "peak end"), ABCA, it is necessary to know the equation of the connecting line, AB. Since the coordinates of these two points have been established, the two point formula is used to calculate the equation of this baseline.

Determination of total area

With the baseline equation established, and the magnitude of all the data points comprising the thermal curve known, integration by Simpson's rule gives the total area of the thermal curve.

Instrument calibration

The total area of the thermal curve of a standard material is equated to its known heat of fusion in order to establish the calorimetric calibration factor for a particular instrument range setting.

Thermal lag arises from the finite response time of the instrument. The maximum slope of the leading edge of the thermal curve of a substance of high thermal conductivity is used as an indication of the thermal lag of the instrument.

Calibration of the temperature scale of the instrument is accomplished with a standard material of known melting point. Indium metal (99.999% pure) is used as the standard material for all three calibration purposes.

Calibration factors for calorimetric sensitivity and thermal lag are retained in a computer subroutine for use in the calculations until replaced with new values.

Evaluation of fraction molten at various temperatures

The ratio, *F*, of the area under the DSC curve up to a given temperature to the total area is taken as the fraction of the sample molten at that temperature. These areas are calculated using the trapezoidal rule.

Corrections

The thermal lag and heat capacity factors are taken into account by the following relationships:

$$T_s = T_p - \frac{V_i S}{60\,L} - \frac{CS^2}{3600\,GL}$$

$$A_c = \left(\frac{CS}{60\,G}\right)^2 \bigg/ L - \frac{V_i CS}{60\,GL}$$

Linearization of 1/F with temperature

Samples less than 100% pure are already molten to a significant extent at the beginning of DSC scans for thermal purity calculation, resulting in serious underestimation of the fraction molten in the early stages if a baseline is constructed as previously described. The consequence is curving of the otherwise linear Van't Hoff plot. Adding the appropriate amount of energy for the underestimation of fraction melted produces a linear $1/F$ vs. T_s plot. The amount is determined by trial to give the straightest line by a least squares test. Provision is included for 30 trials at increments chosen by the operator to give a total correction, k:

$$1/F = \frac{A_T + kA_T}{A_p + a_c + kA_T}$$

Calculation of mole % purity

The mole fraction of impurity is calculated from the slope, Q, of the line that gives the best fit of the corrected data.

$$\chi_2 = Q \times \frac{\Delta H_f}{RT_0^2}$$

The correlation coefficient is calculated as test of the linearity. This coefficient is also used to assess the influence of experimental or computational options.

Options available

In addition to the experimental calorimetric options such as sample weight, temperature range, and scan rate, the operator is provided with the choice of iterative increment (previously mentioned) and the choice of the portion of the DSC curve to be used for the purity calculation. The initial portion of the curve is afflicted by large relative uncertainties in partial area. During the final portion large amounts of energy are being absorbed and then the process halts suddenly; to the extent that the system is not at equilibrium, the Van't Hoff treatment is vitiated. The computer program employs the first half of the thermal curve ($1/F>2$) for the initial purity calculation, but then the operator can select other initial and final points for subsequent calculations.

EXAMPLES OF SYSTEM OPERATION

Experimental data recorded on paper or magnetic tape are transmitted to the time-sharing computer into an intermediate file where it can be edited or from which it is called during calculations.

Data entry

Data are transmitted or calculations performed by the operator through the communications terminal in the conversational mode.

Range calibration or temperature measurement

Analysis of the data in the file requires the appropriate calling sequence. If the data represents a standard material for use in calibrating the calorimeter range setting, the request–response dialogue has the form illustrated in Table 2.

TABLE 2

CALIBRATION INPUT

```
DSC THERMAL PURITY ANALYSIS SYSTEM
TYPE 1 DR 2 --- 1-CALIBRATION   2-ANALYSIS
=1
ENTER SAMPLE IDENTIFICATION
=IN STD
    RANGE        SCAN RATE     SAMPLE WT
   SETTING       (DEG/MIN)     (IN MG)
=5                 1.25          2.588
ENTER HEAT FUSION OF STANDARD IN CAL/GM
=6.79
  INITIAL TEMP TG      DATA RATE        ITERATIVE
    (IN DEG C)       (SEC/DATA AT)      INTERVAL
=153                 1                  .001
LOCATION OF DATA?  1-FILE  2-PAPER TAPE  3-INSTRUMENT
=1
```

TABLE 3

PRINTOUT OF CALIBRATION RESULTS

```
IN STD TEST 2.588 MG R=5 S=1.25 TD=153 CS=2.5 RR=5 SEALED PAN
ORIGINAL N=322
PK SENSE N=45
167 176        5.739        7.011
  SLOPE START       SLOPE END       RANGE VALUE
   156.27083         156.771          0.50874
 THERMAL LAG                         TOTAL AREA        TM
   0.14293                            34.54109       156.249
```

The resulting calibration values are printed out (Table 3) and the range value, *G*, and the thermal lag, *L*, are retained for future reference in Standard File (Table 4).

This portion of the computation is also useful for the objective determination of the extrapolated temperature of the steepest portion of the leading edge of the DSC curve, i.e., the "extrapolated onset" often used along with the peak temperature for the qualitative characterization of samples.

TABLE 4

STANDARD FILE

```
SYSTEM ?EDIT
OLD OR NEW-OLD STAND
READY
-LIST
      1.00000          4.28609          0.08304
      2.00000          0.10346          0.67600
      3.00000          0.50874          0.13764
      4.00000          4.00000          0.
      5.00000          0.50874          0.14293
      6.00000          1.01749          0.28587
      7.00000          1.78113        100.00000
      8.00000          0.50370          1.00000
      9.00000          0.51940          1.22002
     10.00000          0.38562        100.00000
     11.00000          0.25437        100.00000
     12.00000          0.25437        999.00000
     13.00000          0.50874          0.13764
     14.00000         14.00000          0.01000
     15.00000         15.00000          0.01000
     16.00000         16.00000          0.01000
     17.00000         17.00000          0.01000
```

TABLE 5

REQUEST-RESPONSE SEQUENCE FOR PURITY ANALYSIS

```
DSC THERMAL PURITY ANALYSIS SYSTEM
TYPE 1 DR 2 --- 1-CALIBRATION   2-ANALYSIS
=2

ENTER SAMPLE IDENTIFICATION
=BENZOIC ACID

   RANGE        SCAN RATE      SAMPLE WT      MOLEC WT
  SETTING       (DEG/MIN)      (IN MG)        (G/MOLE)
=11                .6              5             122.1

ENTER HEAT CAPACITY CORRECTION IN MCAL/DEG C
=0

  INITIAL TEMP TG      DATA RATE       ITERATIVE
    (IN DEG C)       (SEC/DATA PT)     INTERVAL
=115.725                 2              .001

LOCATION OF DATA? 1-FILE  2-PAPER TAPE  3-INSTRUMENT
=1
```

Heat of fusion and purity determination

For the calculation of purity the initial dialogue is illustrated in Table 5 and the result in Table 6.

Options

The effects of changing the heat capacity correction and range of temperatures considered is detailed in Table 7.

TABLE 6

PRELIMINARY PURITY ANALYSIS

```
ORIGINAL  N = 408
PK SENSE  N = 257
HALF AR   N = 253
```

Benzoic acid

```
X-Y PRINT-OUT DESIRED? --- TYPE YES OR NO
= YES

X-VALUE       Y-VALUES      Y-VALUES     PART AREAS     TEMPERATURE
IN SECS       MV. (ADJ)     NOT ADJ      CORRECTED      (NOT CORR)
426.000       0.087         0.087        11.207         120.005
428.000       0.089         0.089        11.383         120.025
430.000       0.090         0.090        11.562         120.045
432.000       0.092         0.092        11.744         120.065
434.000       0.093         0.093        11.929         120.085
436.000       0.095         0.095        12.118         120.105
438.000       0.097         0.097        12.310         120.125
440.000       0.098         0.098        12.505         120.145
442.000       0.100         0.100        12.703         120.165
444.000       0.102         0.102        12.906         120.185
446.000       0.104         0.104        13.112         120.205
448.000       0.106         0.106        13.321         120.225
450.000       0.108         0.108        13.535         120.245
452.000       0.110         0.110        13.752         120.265
454.000       0.112         0.112        13.974         120.285
456.000       0.114         0.114        14.200         120.305
458.000       0.116         0.116        14.430         120.325
   .            .             .            .               .
   .            .             .            .               .
   .            .             .            .               .
                .                          .

  TOTAL AREA        SUM OF          START         END
THERMAL CURVE       PARTIALS        TEMP          TEMP
673.024             362.470         117.225       122.265

RESTART RUN? --- YES OR NO
= NO

THE 14TH ITERATION PRODUCED THE BEST RESULTS

LINEARIZATION TABLE DESIRED?   YES OR NO
= YES
```

(Table continued on p. 22)

TABLE 6 (*continued*)

```
LINEARIZATION PROCESS OF TEMP (CORR) VS 1/F

PREDICTED        OBSERVED        RESIDUALS        1/FI
117.212           117.224          0.012          72.429
117.222           117.244          0.022          72.284
117.242           117.264          0.022          71.986
117.263           117.284          0.021          71.708
117.283           117.304          0.021          71.419
117.304           117.324          0.020          71.131
   .                 .               .               .
   .                 .               .               .
                                     .

LINEARIZATION PROCESS OF TEMP (CORR) VS 1/F
 HEAT OF FUSION   MOLE FRACT    MOLE %          START      END
  (CORRECTED)      IMPURITY     PURITY          TEMP       TEMP
   4233.162         0.00097    99.90282        117.225    122.262

SLOPE        TEMP TO     ST.ERR SLOPE   CORR COEFF    LAG       RANGE
-0.071       122.373        0.0000         1.000     0.00000   0.25437

  START 1/F            END 1/F
   72.429               1.835
```

TABLE 7

CHANGE OF HEAT CAPACITY OR TEMPERATURE RANGE

```
ENTER 1-NEW PARTIAL AREAS 2-NEW RUN 3-NEW PARAMETERS
      4-NEW LAG.P.CO          5-END OF JOB
=1
ENTER START AND END TEMPERATURES
=121.5      122.1
NEW N=31
X-Y PRINT-OUT DESIRED? --- TYPE YES OR NO
=NO
  TOTAL AREA      SUM OF      START     END
THERMAL CURVE    PARTIALS     TEMP      TEMP
   673.024        160.128    121.505   122.105
RESTART RUN? --- YES OR NO
=NO
THE 14TH ITERATION PRODUCED THE BEST RESULTS
LINEARIZATION TABLE DESIRED?    YES OR NO
=NO
 HEAT OF FUSION   MOLE FRACT    MOLE %          START      END
  (CORRECTED)      IMPURITY     PURITY          TEMP       TEMP
   4239.162         0.00099    99.90070        121.505    122.104
SLOPE        TEMP TO     ST.ERR SLOPE   CORR COEFF    LAG       RANGE
-0.073       122.397        0.0011         1.003     0.00000   0.25437
   START 1/F        END 1/F
```

TESTS OF PROGRAM OPERATION

The validity of various parts of the computer program were subjected to confirmatory tests as a few ambiguities arose. These resulted from an inadequate appreciation of the limitations of the operation of computers ("double precision" is required), or of the need for detail in specifying every aspect of the calculation to the exclusion of other possibilities. The latter point came to focus clearly when checking the ability of the program to detect the peak opening and closing temperatures and the ability to integrate areas.

Temperature detection and integration

The temperature sensing portion of the program was checked by introducing synthetic "data" in the shape of a square pulse. The "temperature" indicated by the program was close to the expected value but the area under the square pulse was indicated to be zero. This was caused by both the peak opening and closing being established in the program as the first point outside the 2σ limit rather than the last point inside the 2σ limit as originally intended. Taking these facts into account, both "temperatures" and the area value expected were obtained exactly in subsequent trials.

Heat of fusion

The combined operations of iteration to compensate for the artificial choice of baseline, and conversion from the experimental millivolt-seconds area to the desired units of calories per mole (with the aid of the previously determined and stored calibration factor) were verified by comparing with manually calculated values. Thus, an area of 431.996 mV-sec. to which 3% is added (by 30 iterations of 0.001) gives 2763.937 cal mol^{-1} for the expected heat of fusion, 2763.939 cal mol^{-1}, for a 5 mg sample of molecular weight 122.1 using an instrument range of 0.25437 mcal sec^{-1} mV^{-1}.

Purity determination

A comprehensive test of the combined operations of the program functions was carried out by introducing values calculated from the Van't Hoff equation and employing the program to determine the purity, heat of fusion and melting point of the pure substance.

The calorimeter ordinate values can be calculated from an equation obtained by differentiating the Van't Hoff equation after substituting $1/F$ by its equivalent, q_0/q_s, i.e., the ratio of the heat required to melt the entire sample to that absorbed at T_s. This gives:

$$\frac{\mathrm{d}q_s}{\mathrm{d}t} = \frac{-RT_0^2 \chi_2 q_0}{\Delta H_f (T_0 - T_s)^2} \frac{\mathrm{d}T_s}{\mathrm{d}t}$$

Neglecting thermal lag (i.e., $T_s = T_p$) and heat capacity corrections, and converting to the previously defined units:

$$V_i = \frac{-RT_0^2 \chi_2 wS}{60\,GM(T_0 - T_p)^2}$$

Ordinate values were calculated for various degrees of purity ($\chi_2 = 0.001$, 0.005, 0.01) at every 0.01 or 0.02°C for several degrees (up to $F = 1$) using the physical constants for benzoic acid: $T_0 = 122.4$°C, $M = 122.1$ g mol^{-1}. The other values employed were $R = 1.987$ cal °C^{-1} mol^{-1}, $w = 5$ mg, $S = 0.6$°C min^{-1}, and $G = 0.25437$ mcal mV^{-1} sec^{-1}. The calculated ordinate values were preceded and followed by a string of zeros to fit the 60-point baseline requirement. A discontinuity at $F = 1$ was taken into account by employing half the value calculated at $F = 1$; this is appropriate as can be demonstrated by geometric analysis of the curve formed by the data points in this region.

TABLE 8

TEST OF PROGRAM WITH THEORETICAL DATA

Data range (°C)	ΔT_s	χ_2 *in*	χ_2*out*	T_0 *out*	ΔH_f *out*	k (%)	*1/F range*
5.1	0.02	0.001	0.00097	122.37	4239	1.4	72–1.8
5.1	0.02	0.005	0.00508	122.42	3990	7.8	13.8–1.8
5.1	0.02	0.01	0.00993	122.38	4062	15	7.6–1.8
4.1	0.01	0.001	0.00098	122.39	4145	1.8	56–1.8
4.1	0.01	0.005	0.00503	122.41	4148	9.2	11.9–1.9
4.1	0.01	0.01	0.0098	122.37	4135	18.0	6.5–1.7

The results (summarized in Table 8) attest to the validity of the program employed for the calculation: the expected values of χ_2 and T_0 are obtained in each case. The correct value of the heat of fusion (4150 cal/mol^{-1}) is obtained using the smaller temperature increment. The coarser increment gives cruder estimates of the area due to the steepness of the curve near $F = 1$, where a single point corresponds to 30% of the area. Thermal lag in real cases rounds the DSC curve, minimizing this effect.

Applications of this system to the determination of purity of pharmaceutical compounds is currently under investigation, but in all cases confirmation by such independent techniques as phase solubility analysis, TLC, etc., as appropriate will be sought.

ACKNOWLEDGEMENTS

The authors are indebted to Mr. Al Chelak for translating physical chemistry into FORTRAN, Dr. John Donahue and Dr. William McSharry for valuable suggestions and to Dr. Jane Sheridan for encouraging these efforts.

REFERENCES

1 *Du Pont Thermal Analysis Applications Brief 900 B35.*
2 *Perkin-Elmer Thermal Analysis Newsletter,* Nos. 5 and 6.