## 4.4 CORRELATION OF THE PHYSICAL PROPERTIES OF ETHANE

### 4.4.1 Concepts Demonstrated

Correlations for heat capacity, vapor pressure, and liquid viscosity for an ideal gas.

### 4.4.2 Numerical Methods Utilized

Polynomial, multiple linear, and nonlinear regression of data with linearization and transformation functions.

### 4.4.3 Excel Options and Functions Demonstrated

Use of the Excel LINEST function for multiple linear and polynomial regression. Use of the Excel Add-In "Solver" for nonlinear regression.

### 4.4.4 Problem Definition

Determine appropriate correlations for heat capacity, vapor pressure, and liquid viscosity of ethane. The data files are given and also the data are available in Appendix F. Compare those correlations with the expressions suggested by the Design Institute for Physical Properties, DIPPR[2].

(a) Compare third-degree and fifth-degree polynomials for the correlation of the heat capacity data (Table A of Appendix F) using both POLY-MATH and Excel by examining the respective variances, confidence intervals, and residual plots.

(b) Use Excel to compare the fifth-degree polynomial for the correlation of the heat capacity data (Table B of Appendix F) with the two DIPPR recommended correlations for the appropriate temperature intervals.

(c) Utilize multiple linear regression in Excel to fit the Wagner equation to the vapor pressure of ethane data found in Table C of Appendix F. Comment on the applicability of the Wagner equation for correlating these data. Compare the correlation obtained by the Wagner equation with that of the Riedel equation recommended by DIPPR.

(d) Use nonlinear regression to fit the Antoine equation to the liquid viscosity data of ethane data found in Table D of Appendix F. Initial estimates of the nonlinear regression parameters should be obtained by linear regression. Verify nonlinear regression results in both POLY-MATH and Excel. Compare the correlation obtained by the Antoine equation with that of the Riedel equation recommended by DIPPR.

### 4.4.5    Solution

This problem can be approached by first setting up the problem in POLYMATH and achieving a solution. Then the problem is exported to Excel from the POLY-MATH program, and the same calculations in Excel are verified between the two software packages. Further use of Excel is emphasized in the detailed problem solution and the generation of the tabular and graphical results.

**(a)** The temperature dependency of the heat capacities of gases is commonly represented by simple polynomials of the form

$$C_p = a_0 + a_1 T + a_2 T^2 + a_3 T^3 + \dots \qquad \textbf{(4-42)}$$

where $C_p$ is the heat capacity in J/kg-mol·K, $T$ is the temperature in K, and $a_0$, $a_1$,... are the coefficients (parameters) of the correlation determined by regression of experimental data. The degree of the polynomial which best represents the experimental data can be determined based on the variance, the correlation coefficient ($R^2$), the confidence intervals of the parameters, and the residual plot. The heat capacity data for ethane gas are given in Appendix F, Tables A and B. There are 19 data points in Table A but they encompass a wider temperature range (1450 K) than the 41 data points in Table B that have a much smaller range of temperature range (400 K).

The data of Table A can be fitted to a third-degree polynomial of the form given by Equation (4-42) by first using the POLYMATH Regression Program. The results of the polynomial obtained with POLYMATH are summarized in Figure 4-26, and the POLYMATH graphical result is given in Figure 4-27. The high value of the correlation coefficient ($R^2 = 0.9971$) as well as the plot of the calcu-

**Model:** Cp = a0 + a1*T_K + a2*T_K^2 + a3*T_K^3

| Variable | Value | 95% confidence |
|---|---|---|
| a0 | 2.505E+04 | 4141.482 |
| a1 | 91.41495 | 24.82297 |
| a2 | 0.0365871 | 0.0378178 |
| a3 | -2.989E-05 | 1.621E-05 |

**General**
Degree of polynomial = 3
Regression including a free parameter
Number of observations = 19

**Statistics**

| R^2 | 0.9975999 |
|---|---|
| R^2adj | 0.9971199 |
| Rmsd | 446.1547 |
| Variance | 4.791E+06 |

**Figure 4–26**   Third-Degree Polynomial Coefficient and Statistics for the Heat Capacity Data of Table A - File **P4-04A.POL**
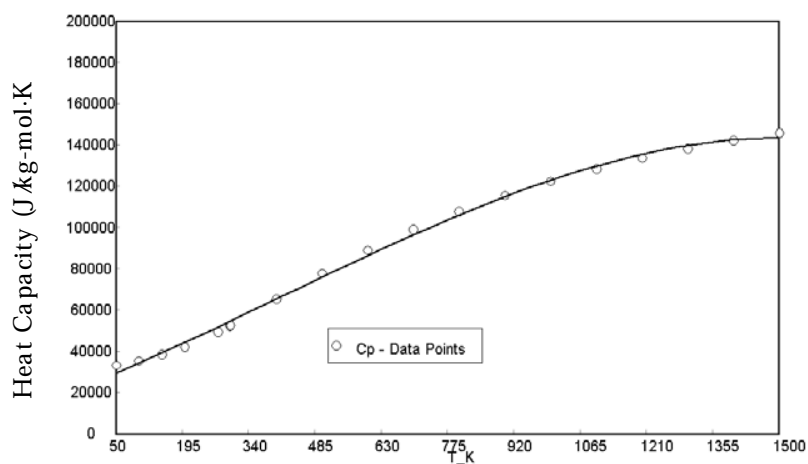
**Figure 4–27**  Third-Degree Polynomial Representation for Heat Capacity of Ethane - File **P4-04A.POL**
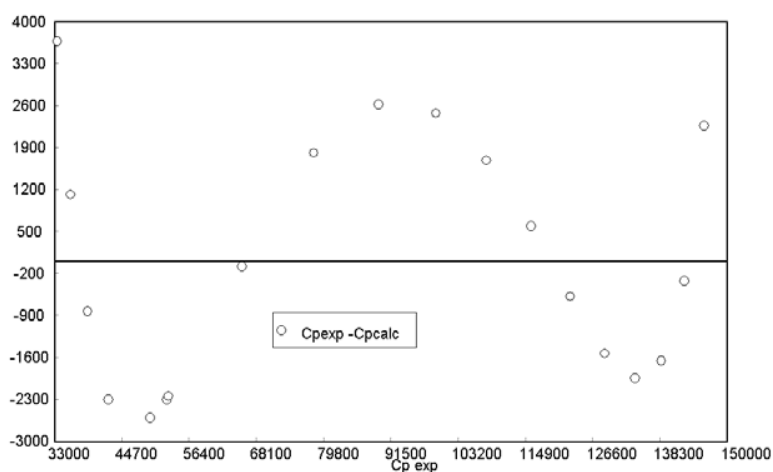


**Figure 4–28**  Residual Plot for Heat Capacity Represented by Third-Degree Polynomial for Data Set A - File **P4-04A.POL**

lated and experimental values seems to indicate that the representation of the data by the third-degree polynomial is quite satisfactory. However, the residual plot of Figure 4–28 shows a clear cyclic pattern, and the error in representation of some of the points is >5% which is well above the common experimental error in heat capacity data. In the case of $a_2$, the confidence interval is slightly larger in absolute value than the parameter itself. Thus the third-degree polynomial representation is unsatisfactory, and better representation should be sought.

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | POLYMATH Polynomial Regression Migration Document | | | | | | | |
| 2 | | | | | | | | |
| 3 | T_K | T_K^2 | T_K^3 | Cp | Cp calc | Cp residual | Cp residual ^2 | |
| 4 | 50 | 2500 | 125000 | 33390 | 29705.72299 | -3684.277013 | 13573897.11 | |
| 5 | 100 | 10000 | 1000000 | 35650 | 34524.72357 | -1125.276432 | 1266247.049 | |
| 6 | 150 | 22500 | 3375000 | 38660 | 39481.83074 | 821.8307351 | 675405.7571 | |
| 7 | 200 | 40000 | 8000000 | 42260 | 44554.63003 | 2294.630034 | 5265326.995 | |
| 8 | 273.16 | 74616.3856 | 20382211.89 | 49540 | 52139.00692 | 2599.006924 | 6754836.99 | |
| 9 | 298.15 | 88893.4225 | 26503573.92 | 52470 | 54762.87925 | 2292.879251 | 5257295.26 | |
| 10 | 300 | 90000 | 27000000 | 52720 | 54957.64721 | 2237.647211 | 5007065.039 | |

**Figure 4–29**  Columns Generated in the Excel Worksheet when a Third-Degree Polynomial Regression is Exported form POLYMATH to Excel - File **P4-04A1.XLS**

| | I | J | K | L | M | N |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | Polynomial Regression of degree 3. Including a free parameter. | | | | | |
| 3 | | a3 | a2 | a1 | a0 | |
| 4 | Coefficients | -3E-05 | 0.036587 | 91.41495 | 25047.24 | |
| 5 | Std.dev.s | 7.61E-06 | 0.017747 | 11.64851 | 1943.445 | |
| 6 | R2, SE (y) | 0.9976 | 2188.737 | #N/A | #N/A | |
| 7 | 95% conf. int. | 1.62E-05 | 0.037818 | 24.82297 | 4141.482 | |
| 8 | Variance | 4790568 | | | | |
| 9 | Sum of Squares | 71858513 | | | | |
| 10 | Model | Cp = a3 * T_K^3 + a2 * T_K^2 + a1 * T_K + a0 | | | | |

**Figure 4–30**  Third-Degree Polynomial Coefficients and Statistics for the Heat Capacity Data of Table A - File **P4-04A1.XLS**

The calculations for the third-degree polynomial can easily be carried out within Excel. This is accomplished from POLYMATH by clicking on the Excel icon from POLYMATH Data Table after the problem is selected for the variable and the desired polynomial degree. Note that an Excel spreadsheet must be open on your computer in order for the "Export to Excel" to take place. The columns generated in the Excel worksheet, after exporting the problem from POLY-MATH, are partially shown in Figure 4–29. The temperature and heat capacity data are found in columns A and D respectively and the formulas for calculating various powers of T are placed in columns B and C. The Excel result is summarized in Figure 4–30 which corresponds very closely to the POLYMATH solution.

In a similar manner, the problem for the fifth-degree polynomial can be setup in POLYMATH and exported to Excel. The resulting worksheet is partially presented in Figure 4–31 where the data columns are shown. The temperature and heat capacity data are found in columns A and F respectively, and the formulas for calculating various powers of T are placed in columns B through E.

Consider now the underlying calculations in the Excel worksheet that are shown in Figure 4–32. The first three rows of this table (cell range L4:Q6) are obtained from Excel's LINEST function. Thus the formula in that range of cells is given by

{=LINEST(F4:F22,A4:E22,TRUE,TRUE)}

where (F4:F22) is the range where the dependent variable, $C_p$ is stored, the second range (A4:E22) is the range where the independent variables (temperature

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | POLYMATH Polynomial Regression Migration Document | | | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | T_K | T_K^2 | T_K^3 | T_K^4 | T_K^5 | Cp | Cp calc | Cp residual | Cp residual ^2 | |
| 4 | 50 | 2500 | 125000 | 6250000 | 312500000 | 33390 | 33762.18787 | 372.1878732 | 138523.8129 | |
| 5 | 100 | 10000 | 1000000 | 100000000 | 10000000000 | 35650 | 35206.23623 | -443.7637733 | 196926.2865 | |
| 6 | 150 | 22500 | 3375000 | 506250000 | 75937500000 | 38660 | 38190.67548 | -469.324517 | 220265.5022 | |
| 7 | 200 | 40000 | 8000000 | 1600000000 | 3.2E+11 | 42260 | 42355.99691 | 95.99690613 | 9215.405987 | |
| 8 | 273.16 | 74616.3856 | 20382211.89 | 5567605000 | 1.52085E+12 | 49540 | 49938.08766 | 398.0876623 | 158473.7869 | |

**Figure 4–31** Fifth-Degree Polynomial Excel Worksheet for the Heat Capacity Data of Table A - File **P4-04A2.XLS**

| | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | Polynomial Regression of degree 5. Including a free parameter. | | | | | | | |
| 3 | | a5 | a4 | a3 | a2 | a1 | a0 | |
| 4 | Coefficients | -7.8E-11 | 3.68E-07 | -0.00065 | 0.481171 | -32.5818 | 34267.42 | |
| 5 | Std.dev.s | 8.82E-12 | 3.4E-08 | 4.71E-05 | 0.028361 | 7.005462 | 544.6506 | |
| 6 | R2, SE (y) | 0.999947 | 348.3187 | #N/A | #N/A | #N/A | #N/A | |
| 7 | 95% conf. int. | 1.9E-11 | 7.34E-08 | 0.000102 | 0.061259 | 15.1318 | 1176.445 | |
| 8 | Variance | 121325.9 | | | | | | |
| 9 | Sum of Squares | 1577237 | | | | | | |
| 10 | Model | Cp = a5 * T_K^5 + a4 * T_K^4 + a3 * T_K^3 + a2 * T_K^2 + a1 * T_K + a0 | | | | | | |

**Figure 4–32** Fifth-Degree Polynomial Coefficients and Statistics for the Heat Capacity Data of Table A - File **P4-04A2.XLS**

and its various powers) are stored. The first logical variable indicates if there is a free parameter (TRUE) in the expression, and the second logical variable indicates whether correlation statistics should be shown (TRUE) in addition to the parameter values.

The regression model parameters are shown in the 4th row of Figure 4–32. The respective parameter standard deviations $\sigma_j$, as provided by the LINEST function, are shown in row 5. The respective 95% confidence intervals are calculated in row 7 by multiplying the $\sigma_j$ by the statistical $t$ distribution value consistent with the number of degrees of freedom (the appropriate $t$ value is inserted by the POLYMATH export utility). The confidence interval of the parameter $a_0$ is calculated, for example, using the formula

```
=2.017*Q5
```

The linear correlation coefficient (R2 = 0.999947) in cell L6 and the standard error on the dependent variable in cell M6 are also calculated by the LINEST function. The Variance is calculated in cell L8 (=(M6)^2), and the Sum of Squares of the Residuals in cell L9 (=SUM(I4:I44)) is calculated from the generated Excel table.

When changes are introduced in the data, the Excel results table (Figure 4–32) will be updated correctly unless there is a change in the number of data points. If the number of data points is reduced or increased, the data range for the LINST function must be changed, and a different $t$ value (reflecting the change in the degrees of freedom) must be introduced.
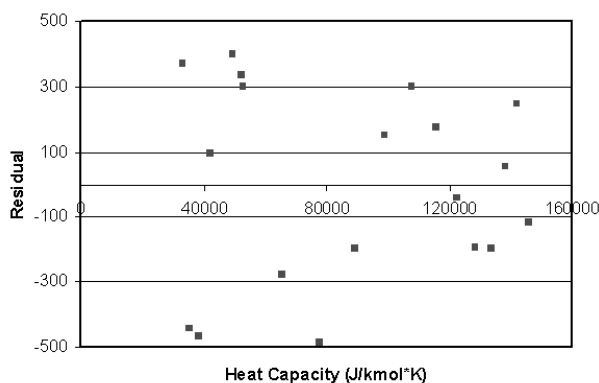
**Figure 4–33** Residual Plot Created in Excel for Heat Capacity Represented by Fifth-Degree Polynomial for the Data Set A

The parameter values for the polynomial shown in Figure 4–32 are used to calculate the "Cp calc" values of Figure 4–31. For example, the formula to calculate "Cp calc" for T = 50 K is

`=$L$4*A4^5+$M$4*A4^4+$N$4*A4^3+$O$4*A4^2+$P$4*A4^1+$Q$4`

Note that these formulas are automatically generated by the POLYMATH software when the export to Excel is requested. The respective residuals, `(Cpcalc-Cp)`, are calculated and placed in column H.

The residual plot, that can be created within Excel, is presented in Figure 4–33. The correlation coefficient is $R^2 = 0.9999$, and the variance has been significantly reduced. All of the confidence intervals are smaller in absolute value than the associated parameter values. The residual plot of Figure 4–33 indicates a random residual distribution with maximum error ~ 1%, which is very similar to the magnitude of the experimental error for this type of data. Thus it can be concluded that the fifth-degree polynomial adequately represents the heat capacity data of Appendix F, Table A.

**(b)** DIPPR[2] recommends an equation for heat capacity of ethane for the temperature range from 200 K through 1500 K

$$C_p = A + B\left[\frac{C/T}{\sinh(C/T)}\right] + D\left[\frac{E/T}{\cosh(E/T)}\right] \qquad \textbf{(4-43)}$$

with parameters A = 4.0326E+04, B = 1.3422E+05, C = 1.6555E+03, D = 7.3223E+04 and E = 7.5287E+02. For the more limited temperature range from 50 K through 200 K, DIPPR recommends using a second-degree polynomial

$$C_p = a_0 + a_1 T + a_2 T^2 \qquad \textbf{(4-44)}$$

with the parameter values $a_0 = 3.1742E+04$, $a_1 = 2.6567E+01$, and $a_2 = 1.2927E-01$.

A comparison of the heat capacity data correlations first requires the deter-

mination of the fifth-order polynomial for the ethane data of Table B in Appendix F. POLYMATH will then be used to obtain the polynomial and subsequently export the problem to Excel for verification of the polynomial representation. The Excel solution will then be modified to carry out the heat capacity calculations using the two DIPPR equations with each applied over the recommended temperature range. A comparison of the polynomial with the DIPPR correlations will then be made in Excel.

The problem can be entered into POLYMATH and the fifth-order polynomial can be used to correlate the data of Table B in the same manner as described in part (a) of this problem. The fifth-degree polynomial problem specified in POLYMATH can then be exported to Excel. The resulting Excel solution is shown in Figure 4–34. It is helpful and good practice to also carry out the POLYMATH polynomial regression in order to verify the Excel solution by comparing the   calculated polynomial coefficients.

The heat capacity values recommended by DIPPR (Equations (4-43) and (4-44)) and the corresponding residual calculations can easily be compared by inserting two new columns in the worksheet immediately to the right of the "Cp residual^2" column I in the Excel worksheet (see Figure 4–36). The five coefficients of Equation (4-43) are entered in the range of cells G48:K48 and the three coefficients of Equation (4-44) are stored in the range of cells G49:I49 as shown in Figure 4–35.

The calculated heat capacity values from the DIPPR equations can be entered in Column J with title "CpD calc" and the residuals are entered in column K with title "CpD residual". The formula for calculating CpD for the first 11 data points $(T \leq 200K)$ is given by the Excel equivalent to Equation (4-44).

=$G$49+$H$49*A4+$I$49*A4^2

Note that this formula refers to $T = 100K$ in Figure 4–36.

| | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | Polynomial Regression of degree 5. Including a free parameter. | | | | | | | |
| 3 | | a5 | a4 | a3 | a2 | a1 | a0 | |
| 4 | Coefficients | 5.07E-09 | -8.2E-06 | 0.004616 | -0.99332 | 147.0626 | 27071.2 | |
| 5 | Std.dev.s | 1.46E-10 | 2.19E-07 | 0.000126 | 0.034084 | 4.342579 | 206.0041 | |
| 6 | R2, SE (y) | 0.999999 | 12.71245 | #N/A | #N/A | #N/A | #N/A | |
| 7 | 95% conf. int. | 2.94E-10 | 4.43E-07 | 0.000254 | 0.068917 | 8.780695 | 416.5403 | |
| 8 | Variance | 161.6065 | | | | | | |
| 9 | Sum of Squares | 5656.227 | | | | | | |
| 10 | Model | Cp = a5 * T_K^5 + a4 * T_K^4 + a3 * T_K^3 + a2 * T_K^2 + a1 * T_K + a0 | | | | | | |

**Figure 4–34** Fifth-Degree Polynomial Coefficients and Statistics from Excel for the Heat Capacity Data of Table B of Appendix F.

| | A or a0 | B or a1 | C or a2 | D | E |
|---|---|---|---|---|---|
| 47 | | | | | |
| 48 | 4.0326E+04 | 1.3422E+05 | 1.6555E+03 | 7.3223E+04 | 7.5287E+02 |
| 49 | 3.1742E+04 | 2.6567E+01 | 1.2927E-01 | | |

**Figure 4–35** Coefficients of the DIPPR Equations - File **P4-04.XLS (Cp_Table B)**

|   | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **POLYMATH Polynomial Regression Migration Document** | | | | | | | | | | |
| 2 | | | | | | | | | | | |
| 3 | T_K | T_K^2 | T_K^3 | T_K^4 | T_K^5 | Cp | Cp calc | Cp residual | Cp residual ^2 | CpD calc | CpD resid |
| 4 | 100 | 10000 | 1000000 | 1E+08 | 1E+10 | 35698 | 35695 | -3.0314659 | 9.18978525 | 3.57E+04 | -6.6 |
| 5 | 110 | 12100 | 1331000 | 1.46E+08 | 1.61E+10 | 36249 | 36260 | 10.8427755 | 117.565781 | 3.62E+04 | -20.463 |
| 6 | 120 | 14400 | 1728000 | 2.07E+08 | 2.49E+10 | 36817 | 36826 | 8.65949312 | 74.98682118 | 3.68E+04 | -25.472 |
| 7 | 130 | 16900 | 2197000 | 2.86E+08 | 3.71E+10 | 37401 | 37402 | 0.63596182 | 0.404447442 | 3.74E+04 | -20.627 |
| 8 | 140 | 19600 | 2744000 | 3.84E+08 | 5.38E+10 | 38003 | 37996 | -7.2374236 | 52.38030027 | 3.80E+04 | -7.928 |
| 9 | 150 | 22500 | 3375000 | 5.06E+08 | 7.59E+10 | 38628 | 38615 | -13.136289 | 172.5620765 | 3.86E+04 | 7.625 |
| 10 | 160 | 25600 | 4096000 | 6.55E+08 | 1.05E+11 | 39279 | 39265 | -14.341419 | 205.6762906 | 3.93E+04 | 23.032 |
| 11 | 170 | 28900 | 4913000 | 8.35E+08 | 1.42E+11 | 39961 | 39950 | -11.1779 | 124.945454 | 4.00E+04 | 33.293 |
| 12 | 180 | 32400 | 5832000 | 1.05E+09 | 1.89E+11 | 40680 | 40674 | -5.9542598 | 35.4532098 | 4.07E+04 | 32.408 |
| 13 | 190 | 36100 | 6859000 | 1.3E+09 | 2.48E+11 | 41439 | 41440 | 1.0983953 | 1.206472241 | 4.15E+04 | 17.377 |
| 14 | 200 | 40000 | 8000000 | 1.6E+09 | 3.2E+11 | 42243 | 42250 | 6.88723704 | 47.43403401 | 4.22E+04 | -16.8 |
| 15 | 210 | 44100 | 9261000 | 1.94E+09 | 4.08E+11 | 43092 | 43105 | 12.5185765 | 156.7147568 | 43221.87 | 129.86679 |
| 16 | 220 | 48400 | 1.1E+07 | 2.34E+09 | 5.15E+11 | 43989 | 44004 | 15.3587237 | 235.8903934 | 43981.94 | -7.057476 |
| 17 | 230 | 52900 | 1.2E+07 | 2.8E+09 | 6.44E+11 | 44934 | 44949 | 15.0948477 | 227.8544275 | 44831.64 | -102.3636 |
| 18 | 240 | 57600 | 1.4E+07 | 3.32E+09 | 7.96E+11 | 45924 | 45938 | 13.7958363 | 190.3251001 | 45764.11 | -159.8942 |
| 19 | 250 | 62500 | 1.6E+07 | 3.91E+09 | 9.77E+11 | 46959 | 46969 | 9.97315603 | 99.46384127 | 46771.34 | -187.6573 |
| 20 | 260 | 67600 | 1.8E+07 | 4.57E+09 | 1.19E+12 | 48036 | 48041 | 4.64171187 | 21.54548908 | 47844.67 | -191.335 |

**Figure 4–36** Addition of DIPPR Equation Calculations to Excel Spreadsheet - File **P4-04.XLS (Cp_Table B)**

The remaining data points use the Excel equivalent to Equation (4-43) as it is applied to temperatures greater than 200 K. This is shown below for cell H19 in Figure 4–36.

```
=$G$48+$H$48*(($I$48/A19)/SINH($I$48/
    A19))^2+$J$48*(($K$48/A19)/COSH($K$48/A19))^2
```

The residuals for the DIPPR equations are calculated in Column K by entering the formula for the difference between the DIPPR result in Column J and the measured $C_p$ in Column F.

The residuals of the heat capacity values calculated by fifth order polynomial in Column H and the DIPPR equations in Column K can be plotted in Excel as shown in Figure 4–37. The maximal error in polynomial representation is < 0.1% and the maximal error in the DIPPR correlation is about 0.5%. Note that
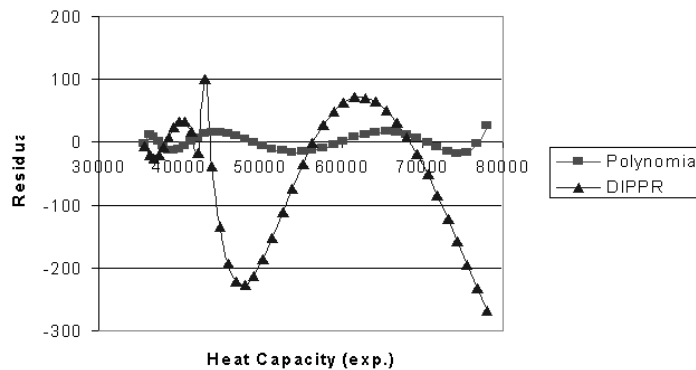


**Figure 4–37** Residual Comparison of Heat Capacity Representation by a Fifth-Degree Polynomial and the DIPPR Equations for Data Set B - File **P4-04.XLS**

the larger error for DIPPR is expected as the DIPPR correlation of Equation (4-43) is for a much larger range of temperature. The residuals of both correlations show cyclic trends, and these trends can probably be attributed to prior smoothening of the experimental data.

(c) The Wagner equation is considered by many as the most appropriate model to represent the vapor pressure data over the full range between the triple point and critical point. The most widely used form of the Wagner equation is

$$\ln P_R = \frac{a\tau + b\tau^{1.5} + c\tau^3 + d\tau^6}{T_R} \tag{4-45}$$

where $T_R = T/T_C$ is the reduced temperature, $P_R = P/P_C$ is the reduced pressure, and $\tau = 1 - T_R$. For ethane, $T_C = 305.32$ K, $P_C = 4.8720E+06$ Pa and the triple point temperature is 90.352 K. Thus the data in Table C of Appendix F cover almost the full range between the triple point and the critical point, and the Wagner equation is appropriate for correlation of these data.

The use of Excel for solving this problem is preceded by the use of POLY-MATH to enter the data into the POLYMATH Data Table. The ability to easily transform data is utilized in POLYMATH to define additional columns in the Data Table as transformation functions defined by

```
TR = T / 305.32
lnPR = ln(P/4872000)
t = (1-TR)/TR
t15 = (1-TR)^1.5/TR
t3 = (1-TR)^3/TR
t6 = (1-TR)^6/TR
```

The resulting POLYMATH Data Table is partially shown in Figure 4–38.

These data transformations allow Multiple Linear Regression to fit the data to the Wagner equation with lnPr as the dependent variable and the independent variables t, t15, t3, and t6. Note that in this Multiple Linear Regression there should be no free parameter; thus, the POLYMATH Data Table option "through origin" should be marked. This problem is exported to Excel after it is setup in the POLYMATH Regression Data Table.

| | T_K | P_Pa | Tr | lnPr | t | t15 | t3 | t6 |
|---|---|---|---|---|---|---|---|---|
| 01 | 92 | 1.7 | 0.3013232 | -14.86839 | 2.318696 | 1.938126 | 1.13187 | 0.3860338 |
| 02 | 94 | 2.8 | 0.3078737 | -14.3694 | 2.248085 | 1.870275 | 1.07692 | 0.3570586 |
| 03 | 96 | 4.6 | 0.3144242 | -13.87296 | 2.180417 | 1.805374 | 1.024827 | 0.3302303 |
| 04 | 98 | 7.2 | 0.3209747 | -13.42493 | 2.11551 | 1.743244 | 0.9754096 | 0.305383 |
| 05 | 100 | 11 | 0.3275252 | -13.00112 | 2.0532 | 1.683718 | 0.9285029 | 0.2823653 |

Figure 4–38 POLYMATH Data Table with Original and Transformed Data Columns - File P4-04C.POL

| | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|
| 1 | **Migration Document** | | | | | | | | |
| 2 | | | | | Multiple Linear Regression. No free parameter. | | | | |
| 3 | lnPr calc | lnPr residual | lnPr residual ^2 | | | a4 | a3 | a2 | a1 |
| 4 | -14.8538 | 0.014542267 | 0.000211478 | | **Coefficients** | -1.2599194 | -1.67115626 | 1.28954942 | -6.4584809 |
| 5 | -14.357 | 0.012431298 | 0.000154537 | | **Std.dev.s** | 0.14859015 | 0.135235971 | 0.1086674 | 0.04802602 |
| 6 | -13.8828 | -0.009812224 | 9.62797E-05 | | **R2, SE (y)** | 0.99999741 | 0.009668808 | #N/A | #N/A |
| 7 | -13.4298 | -0.004871516 | 2.37317E-05 | | **95% conf. int.** | 0.29123669 | 0.265062504 | 0.21298811 | 0.094131 |
| 8 | -12.9967 | 0.004373583 | 1.91282E-05 | | **Variance** | 9.3486E-05 | | | |
| 9 | -12.5824 | -0.016578932 | 0.000274861 | | **Sum of Squares** | 0.00962904 | | | |
| 10 | -12.1856 | -0.005438393 | 2.95761E-05 | | **Model** | lnPr = a4 * t + a3 * t15 + a2 * t3 + a1 * t6 | | | |

**Figure 4–39** Wagner Equation Model Results for the Ethane Vapor Pressure - File **P4-04.XLS** (Vp_Regress)
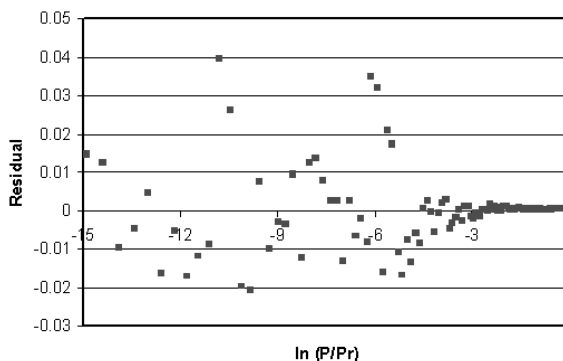


**Figure 4–40** Residual Plot in Excel for Ethane Vapor Pressure Data Represented by the Wagner Equation - File **P4-04.XLS** (Vp_Regress)

The Excel results after export from POLYMATH for fitting the Wagner equation to the vapor pressure data are partially presented in Figure 4–39, and the residuals are plotted in Figure 4–40. The correlation coefficient is $R^2$ = 0.99999, and all the confidence intervals are smaller in absolute value than the associated parameter values. The residual plot shows random residual distribution, and the maximum error is < 1% which is very similar to the magnitude of the experimental error for this type of data. Thus it can be concluded that the Wagner equation adequately correlates the vapor pressure of ethane over the experimental temperature range.

The Riedel equation recommended by DIPPR for vapor pressure data of ethane is given by

$$\ln P = A - \frac{B}{T} + C \ln T + D T^E \qquad \text{(4-46)}$$

with the parameters $A$ = 51.857, $B$ = -2600, $C$ = -5.13, $D$ = 1.49E-05 and $E$ = 2.

The comparison between the Wagner equation and the Riedel equation can be carried out by creating new Excel worksheet which utilizes the Wagner equa-

tion variables and results given in the POLYMATH to Excel worksheet shown in Figure 4–39. Some of the information entered in this prepared worksheet is shown in Figure 4–41 (only four rows of data, out of the 107 data points in this case, are shown). The measured temperature and vapor pressure data are inserted in columns A and B.

| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | **A** | **B** | **C** | **D** | **E** | | | | |
| 2 | | | 5.1857E+01 | -2.60E+03 | -5.13E+00 | 1.49E-05 | 2.00E+00 | | | | |
| 3 | Temp (K) | Vapor Pressure (Pa) | lnPr | lnPr calc | lnPr residual | lnPr Calc DIPPR | lnPr Res DIPPR | Calc VP Wagner | VP Res Wagner | Calc VP DIPPR | VP Res DIPPR |
| 4 | 92 | 1.7 | -14.86839 | -14.8538477 | 0.014542267 | -14.8516 | 0.01677 | 1.7249 | 0.0249 | 1.72875 | 0.0287 |
| 5 | 94 | 2.8 | -14.3694 | -14.3569687 | 0.012431298 | -14.3554 | 0.01403 | 2.83501 | 0.03501 | 2.83956 | 0.0396 |
| 6 | 96 | 4.6 | -13.87296 | -13.8827722 | -0.00981222 | -13.8817 | -0.00876 | 4.55508 | -0.0449 | 4.5599 | -0.0401 |
| 7 | 98 | 7.2 | -13.42493 | -13.4298015 | -0.00487152 | -13.4292 | -0.0043 | 7.16504 | -0.035 | 7.16916 | -0.0308 |

**Figure 4–41** Worksheet for Comparison of Vapor Pressure Correlation by Wagner and DIPPR Equations - File **P4-04.XLS** (Vp_Compare)

The data of "lnPr" and "lnPr calc" (columns C and D in Figure 4–41) are copied from the POLYMATH migration worksheet that is partially shown in Figure 4–39. Note that in order to paste the "lnPr calc" values, the "Paste Special" "Values" should be used otherwise error messages will be obtained (and the data columns and the coefficients of the Wagner equation will not be copied into the new worksheet).

In the 2nd row, the numerical values of the Riedel equation parameters are entered with their names shown in the 1st row. In column E, the "lnPr Calc DIPPR" is calculated using the DIPPR recommended equation by manually entering the formula for cell D4.

= ($C$2+$D$2/A4+$E$2*LN(A4)+$F$2*(A4)^$G$2)-LN(4872000)

Then this formula is copied to all the cells below for the entire data set.

The residual plot of the "lnPr Res DIPPR" in this case is very similar to the residual plot obtained for the Wagner equation (Figure 4–40). The comparison between the two equations is more meaningful if it is carried out with the help of the residual plots based on the pressure (instead of $\ln(P_R)$). The preparation of such a plot is left as an exercise for the reader.

**(d)** A recommended correlation for viscosity of liquids by Perry[4] is similar to the Antoine equation for vapor pressure and given by

$$\ln\mu = A + \frac{B}{T+C} \tag{4-47}$$

where $\mu$ is the viscosity and the parameters are $A$, $B$, and $C$. If $T$ is expressed in degrees K, then parameter $C$ can be approximated by $C = 17.71 - 0.19T_b$ where $T_b$ is the normal boiling point in K. For ethane, the normal boiling point is 184.55 K, and thus the approximate value of $C$ is -17.35.

Equation (4-47) is nonlinear and can be fitted to the experimental viscosity data of Table D in Appendix F using general nonlinear regression. However good initial estimates are necessary for the nonlinear regression. These can be obtained by linearizing Equation (4-47) using the approximate value of $C$ for

ethane to obtain

$$\ln\mu \;=\; A + \frac{B}{T-17.35} \;=\; a_0 + a_1\!\left(\frac{1}{T-17.35}\right) \;=\; a_0 + a_1 X_1 \;\; \text{or} \;\; Y = a_0 + a_1 X_1 \qquad \textbf{(4-48)}$$

Thus, the linear form can be used in the POLYMATH Data Table containing the viscosity and temperature data by creating additional columns to calculate the transformed variables $Y = \ln\mu$ and $X_1 = 1/(T-17.35)$. A portion of the POLYMATH Data Table which utilizes these transformed variables and is set up for the linear regression of Equation (4-48) is shown in Figure 4–42. The results of the POLYMATH Linear Regression are shown in Figure 4–43. These results provide the initial estimates of $A = -11.1$, $B = 364.6$ and $C = -17.35$ for the nonlinear regression of Equation (4-47).
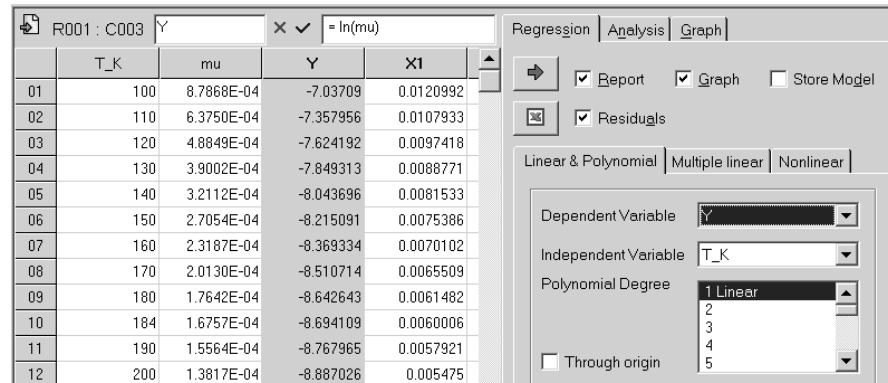


**Figure 4–42** Setup of POLYMATH Linear Regression for Equation (4-48) - File **P4-04D1.POL**

**Model:** Y = a0 + a1*X1

| Variable | Value | 95% confidence |
|---|---|---|
| a0 | -11.09516 | 0.2912278 |
| a1 | 364.5876 | 44.46516 |

**Figure 4–43** Linear Regression Results from POLYMATH for Equation (4-48) - File **P4-04D1.POL**

The nonlinear regression can be set up in POLYMATH and then exported to Excel. The setup of the POLYMATH Nonlinear Regression is shown in Figure 4–44 which gives the results that are summarized in Figure 4–45 where some 73 iterations were required.



**Figure 4–44** Nonlinear Regression Setup in POLYMATH for Equation (4-48) - File **P4-04D1.POL**

**Model:** $Y = A + B/(T\_K + C)$

| Variable | Initial guess | Value | 95% confidence |
|----------|---------------|-------|----------------|
| A | -11.1 | -18.60756 | 5.789612 |
| B | 364.6 | 6568.593 | 7779.302 |
| C | -17.3 | 476.3679 | 395.2097 |

**Nonlinear regression settings**
Max # iterations = 128

**Precision**

| | |
|---|---|
| R^2 | 0.9886379 |
| R^2adj | 0.9875016 |
| Rmsd | 0.0199363 |
| Variance | 0.0105127 |

**General**

| | |
|---|---|
| Sample size | 23 |
| Model vars | 3 |
| Indep vars | 1 |
| Iterations | 73 |

**Figure 4–45** Nonlinear Regression Result in POLYMATH for Equation (4-48) - File **P4-04D1.POL**

The export of the POLYMATH setup for Nonlinear Regression to Excel by pressing the Excel icon gives the initial worksheet that is partially shown in Figure 4–46. Note that this problem in Excel must be solved by using the Excel Add-

### Nonlinear Regression Migration Document

| | Y residual | Y residual ^2 | (Y - Yavg)^2 | (Ycalc - Yavg)^2 | | Nonlinear Regression | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | | | | | | A | B | C |
| 626739 | 0.348463261 | 0.121426644 | 3.376788799 | 4.77889047 | Coefficients | | -11.1 | 364.6 | -17.35 |
| 759849 | 0.193196151 | 0.037324753 | 2.300494439 | 2.923875115 | R2, SE (y) | | 0.932702 | 0.232811 | |
| 124696 | 0.076067304 | 0.005786235 | 1.563755482 | 1.759786363 | Variance | | 0.054201 | | |
| 426542 | -0.014113542 | 0.000199192 | 1.051406384 | 1.022662057 | Average Y | | -8.87469 | | |
| 313494 | -0.083617494 | 0.006991885 | 0.690557829 | 0.558577759 | Model | | Y = A+B/(T_K+C) | | |
| 412494 | 0.126222494 | 0.019592922 | 0.435076239 | 0.273822584 | | | | | |

**Figure 4–46**  Nonlinear Regression Exported to Excel - Initial Worksheet

In called "Solver." This Add-In should be available from the drop-down menu in Excel under "Tools" and then "Add-Ins..."

The objective function for the nonlinear regression problem within Excel is the sum of squares of the Y residuals that is found in the cell at the base of the "Y residual ^2" column.

When Solver is called from the "Tools" menu in Excel to perform the nonlinear regression, an interface appears in which the "Solver Parameters" must be entered. Solver requires that the Target Cell be set as the sum of squares of the Y residuals which should be minimized. Also the Coefficients cells for A, B, and C must be identified in the "By Changing Cells" entry box. This is shown in Figure 4–47. In the "Equal To:" field of the Solver it is important to move the marking to Min (from the default Max marking). After a mouse click on the "Solve" button,
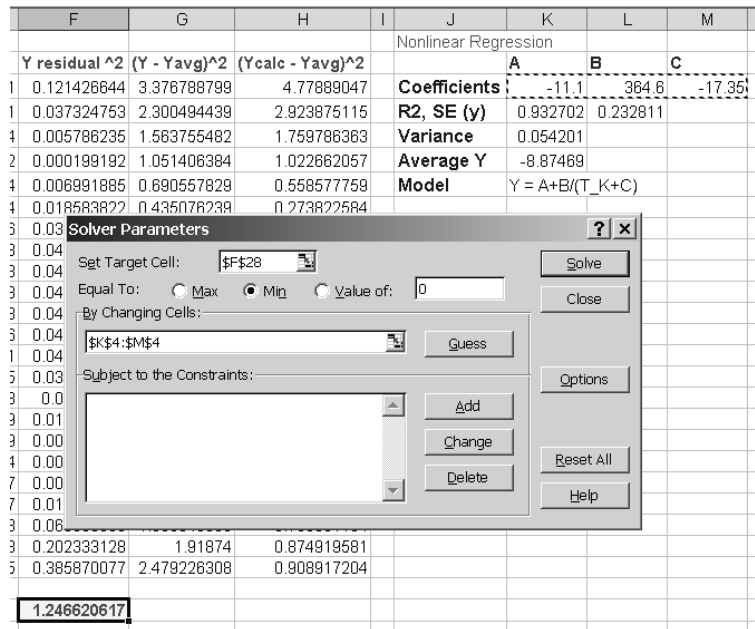


**Figure 4–47**  Use of the Excel Solver Add-In for Nonlinear Regression

the Coefficients are changed to the converged values. In this Solver solution shown in Figure 4–48, the results are similar to the POLYMATH Nonlinear Regression parameters as summarized in Figure 4–45. Note the convergence of POLYMATH and the Excel Solver Add-In are very dependent upon the initial estimates and the particular numerical method that is used. For this problem in POLYMATH, the L-M algorithm has been used, and the number of iterations needed to be increased from the default value. Other algorithms may give different results.

The residual plot from Excel reproduced in Figure 4–49 has a cyclic pattern and considerable errors. This indicates that this model for correlation of ethane viscosity is not very satisfactory. Many more models do exist which could be fitted to these ethane data.

| J | K | L | M |
|---|---|---|---|
| Nonlinear Regression | | | |
| | A | B | C |
| Coefficients | -18.5391 | 6476.254 | 471.6516 |
| R2, SE (y) | 0.988663 | 0.095612 | |
| Variance | 0.009142 | | |
| Average Y | -8.87016 | | |
| Model | Y = A+B/(T_K+C) | | |

**Figure 4–48**  Solver Results for the Excel Nonlinear Regression of Equation (4-47) - File **P4-04.XLS** (Antoine (2))
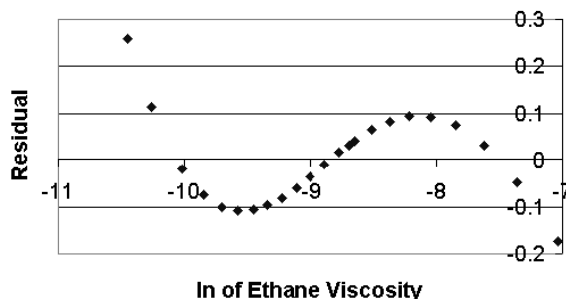


**Figure 4–49**  Residual Plot from Excel for Equation (4-47) - File **P4-04.XLS** (Antoine

The comparison with the Riedel equation using the parameters recommended by DIPPR follows the same procedure that was followed in connection with the vapor pressure data correlation and discussed in the solution to the previous part (c).

The Antoine and Riedel equation representations of the liquid viscosity are compared in Figure 4–50. The residuals of the Riedel equation seem to follow a cyclic pattern as do the residuals of the Antoine equation but the errors are considerably smaller.
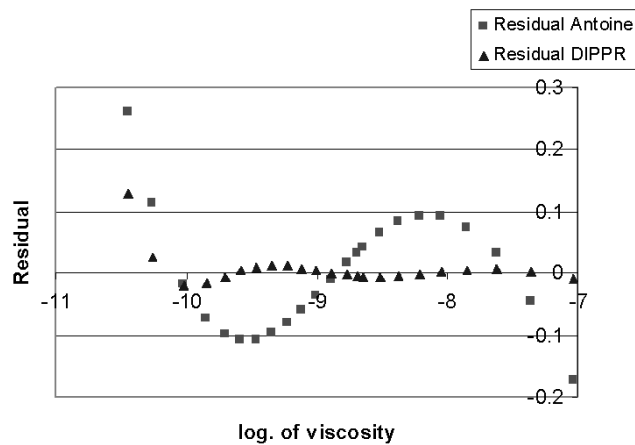


**Figure 4–50**  Comparison of Viscosity Represented by the Antoine Equation (4-47) and th Equation with the DIPPR Recommended Constants - File **P4-04.XLS** (Anto

The problem solution files are found in directory CHAPTER 4 and designated **P4-04A.POL**, **P4-04B.POL**, **P4-04C.POL**, **P4-04D1.POL**, **P4-04D2.POL**, and **P4-04.XLS**.