

# COMMON ERRORS IN NUMERICAL PROBLEM SOLVING – HOW CAN THEY BE DETECTED AND PREVENTED

*Mordechai Shacham, Ben-Gurion University, Beer-Sheva, Israel*

*Michael B. Cutlip, University of Connecticut, Storrs, CT 06269, USA*

*Michael Elly, Intel Corporation, Qiryat Gat, Israel*

## Introduction

The introduction of mathematical software packages such as Excel<sup>1</sup>, MATLAB<sup>2</sup> and Polymath<sup>3</sup> into the field of engineering problem solving has brought the benefits of higher speed and precision in obtaining the results. However, numerical solution techniques have also introduced some new sources of errors. These errors may often pass undetected because of the lack of experience in the use of the new computational tools. A few examples of errors originating from numerical sources follow.

### 1. *Regression and Analysis of Data*

The most common errors in regression of data originate from the use of regression models with too many or too few parameters<sup>[1]</sup>, fitting polynomial models without proper scaling (standardization) of the temperature data, correlation of data when the model equations are improperly linearized<sup>[1]</sup> or when experimental design for obtaining the data is not satisfactory<sup>[3]</sup>.

### 2. *Systems of Nonlinear Algebraic Equations (NLE)*

There are many examples in the literature showing that identification of good initial guesses and formulating the problem correctly for enabling convergence to a solution of an NLE system may represent a major challenge<sup>[4]</sup>. Even if the solution is found, it may be infeasible in the physical sense (solution with negative concentrations, for example) or a false solution caused by improper variable and/or function scaling.

### 3. *Ordinary Differential Equations (ODE)*

Indiscriminate use of default error tolerances of the ODE solver tools is the most common source of errors in solving ordinary differential equations<sup>[5]</sup>. Failure to use the proper integration algorithm (stiff vs. non-stiff), carelessly rounding numbers in the model equations, using the model outside the domain of its validity and the use of inadequate resolution in presenting the results have been documented<sup>[1],[2]</sup> as additional common sources of errors in solving ODEs.

Being aware of potential causes to errors in numerical problem solving is a preliminary requirement to detection and prevention of such errors. In this paper three potential causes of errors are described and demonstrated. Means to prevent these errors are proposed.

## No Consideration Given to Problem Parameter Uncertainties

The importance of the analysis of the sensitivity of the problem solution to small changes in the parameter values has been recognized in the field of optimization. Edgar et al. (2001), for example, wrote "You should always be aware of the sensitivity of the optimal answer ... Parameter values usually contain errors or uncertainties. Information concerning the sensitivity of the optimum to changes or variations in a parameter is therefore very important in optimal process design." However, sensitivity analysis is rarely carried for other types of problems. In

---

<sup>1</sup> Excel is a trademark of Microsoft Corporation (<http://www.microsoft.com>)

<sup>2</sup> MATLAB is a trademark of The Math Works, Inc. (<http://www.mathworks.com>)

<sup>3</sup> POLYMATH is a product of Polymath Software (<http://www.polymath-software.com>)

this section, the importance of the sensitivity analysis in solving systems of nonlinear algebraic equations will be demonstrated.

The example considered was originally published by Froment and Bishoff (1990), and its solution using Mathcad was presented by Parulekar (2006). The reaction considered is the catalytic hydrogenation of olefins (component A) in an isothermal CSTR. Differential mass balance on component A yields

$$\frac{dC_A}{dt} = \frac{C_{A0} - C_A}{\tau} - r_A \quad (1)$$

where  $C_{A0}$  is the inlet concentration of A (mol/L),  $\tau = V / v_0$  is the space time (s),  $V$  is the reactor volume (L),  $v_0$  is the inlet flow rate (L/s) and  $r_A$  is the reaction rate per unit reactor volume (mol/L·s):

$$r_A = \frac{C_A}{(1 - C_A)^2} \quad (2)$$

The concentration of A at steady-state is sought for the operating conditions:  $C_{A0} = 13$  mol/L;  $V = 10$  L; and  $v_0 = 0.2$  L/s. Finding the steady-state solution requires solving the nonlinear algebraic equation  $f(C_A) = 0$  where

$$f(C_A) = \frac{C_{A0} - C_A}{\tau} - r_A \quad (3)$$

In Figure 1,  $f(C_A)$  is plotted versus  $C_A$  for the region  $0 \leq C_A \leq 10$ . Observe that in this region there are three real solutions. The solutions as obtained by the Polymath 6.1 software package are:  $C_A^* = [0.7515358, 2.130933, \text{ and } 8.117532]$ . Polymath displays the results showing 7 significant digits. Mathcad displays the same results with 20 significant digits (Parulekar, 2006). The discrepancy between the number of significant digits in the CSTR parameters (one or two) and the number of those digits reported in the solution (up to twenty) cannot be missed. The determination of the number of significant digits to be shown in the documentation of the problem solution should be based, in this particular case, on the uncertainty associated with the parameter values.

In Table 1, the values of  $C_A$  at the solution are presented for various levels of uncertainty in the flow rate  $v_0$ . Assuming an uncertainty of  $10^{-4}$  L/s (thus flow rate of 0.2001 L/s instead of 0.2 L/s) yields solution values which match the base solution only in the first two digits. For uncertainty of  $10^{-3}$  L/s there is only one correct digit in the first two roots and two correct digits in the largest root. For an uncertainty of 0.01 L/s there is not even a single correct digit in the first two roots and one in the largest root. Thus, any realistic estimate for the uncertainty in the value of the flow rate leads to the conclusion that the solutions for  $C_A$  should be rounded up to two significant digits as all the additional digits are uncertain. It should be emphasized that this analysis applies only to physical quantities which obtained as the final results of the computation. There are other cases where rounding the numbers may lead to significant errors. This will be demonstrated in one of the following examples.

As the value of the flow rate is specified by one significant digit, only one possible assumption regarding the uncertainty could be that the flow rate was measured (or controlled)

with such an accuracy, thus its value could vary in the range of  $0.15 \text{ L/s} \leq v_0 \leq 0.24 \text{ L/s}$ . In the last two rows of Table 1 the solutions for the two extreme values of  $v_0$  are shown. For  $v_0 = 0.24 \text{ L/s}$  there is only one real root at  $C_A = 9.36922 \text{ mol/L}$  while for  $v_0 = 0.15 \text{ L/s}$  there is one real root at  $C_A = 0.3419 \text{ mol/L}$ . Thus, in these cases there is only one root instead of the three roots for the base value of  $v_0 = 0.2 \text{ L/s}$ , and the roots are completely different.

### Use of Inappropriate Statistics and Graphical Representations in Analysis of Regression Models

The use of inappropriate statistics and graphical representation may often lead to the acceptance unsatisfactory regression models. This phenomenon will be demonstrated by fitting the Clausius-Clapeyron equation to vapor pressure ( $P_V$ ) data of ethane. The vapor pressure data, containing  $N = 107$  points, are available<sup>[8]</sup> in the temperature range of  $T = 92 \text{ K}$  ( $P_V = 1.7 \text{ Pa}$ ) through  $T = 304 \text{ K}$  ( $P_V = 4.738 \cdot 10^6 \text{ Pa}$ ). This range covers essentially the liquid phase region between the triple point temperature ( $= 90.352 \text{ K}$ ) and the critical temperature ( $T_C = 305.32 \text{ K}$ ) of ethane. The Clausius-Clapeyron equation is considered inappropriate for modeling vapor pressure for such a wide range of temperatures and we want to examine what kind of statistics and/or graphical representation can reveal this inappropriateness.

The Clausius-Clapeyron equation is given by

$$\ln(P_V) = A + \frac{B}{T} \quad (4)$$

where A and B are coefficients calculated by regression of experimental data.

In Figure 2  $\ln(P_V)_{\text{exp}}$  is plotted versus  $1/T$ . The straight line fitted to the data is also shown. The linear equation obtained is

$$\hat{y} = 21.816 + 1919.5x \quad (5)$$

where  $x = 1/T$  and  $\hat{y} = \ln(P_V)_{\text{calc}}$ . The correlation coefficient  $R^2$  is defined

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (6)$$

where  $\bar{y}$  is the sample mean of the dependent variable. The value of R is bounded:  $0 \leq R \leq 1$ . If R is close to 1, there is a strong correlation between the variables, whereas a value close to zero indicates a weak or no correlation. The  $R^2$  value in this case (as shown in Figure 2) is 0.999, very close to 1. Thus, the plot of  $\ln(P_V)_{\text{exp}}$  versus  $1/T$  and the statistical  $R^2$  both indicate that the Clausius-Clapeyron equation represents the vapor pressure data adequately over the entire range from the triple point to  $T_C$ .

Another widely used graphical representation for checking adequacy of a correlation is the plot of the calculated value of the dependent variable  $\hat{y}$  versus the experimental value  $y$ . In case of a good fit a straight line should be obtained with the slope  $\sim 1.0$ . In Figure 3,  $\ln(P_V)_{\text{calc}}$  (as calculated by the Clausius-Clapeyron equation) is plotted versus  $\ln(P_V)_{\text{exp}}$ . Observe that the data points are aligned along a straight line with the slope of 0.9999 with only small deviations. Thus, this type of representation indicates also a good fit between the Clausius-Clapeyron equation and the vapor pressure data over the entire range. However, comparing  $\ln(P_V)_{\text{calc}}$  and  $\ln(P_V)_{\text{exp}}$  at  $T = 92 \text{ K}$ , for example, yields  $\ln(P_V)_{\text{exp}} = 0.53$ ;  $\ln(P_V)_{\text{calc}} = 0.95$ , which is an  $\sim 80 \%$

difference. This definitely cannot be defined as adequate representation of the experimental data.

The proper graphical means to test the goodness of fit between a regression model and the data is the "residual" plot. The residual  $\hat{\varepsilon}_i$  is defined as

$$\hat{\varepsilon}_i = y_i - \hat{y}_i \quad (7)$$

A random distribution of the residuals around zero indicates that the regression model represents the data correctly. A definite trend or pattern in the residual plot may indicate either lack of fit of the model or that the assumed error distribution for the data (i.e. random error distribution in  $y$ ) is incorrect. In Figure 4 the residual:  $\ln(P_V)_{\text{exp}} - \ln(P_V)_{\text{calc}}$  is plotted versus  $\ln(P_V)_{\text{exp}}$ . Observe that there is definite trend (curvature) in the residuals, indicating that the Clausius-Clapeyron equation cannot represent the data adequately.

### Careless Rounding of Model Parameters.

In the DIPPR database<sup>[9]</sup>, the Riedel equation is recommended for modeling vapor pressure data. The Riedel equation can be written

$$\ln(P_V) = A + \frac{B}{T} + C \ln T + DT^E \quad (8)$$

where  $A, B, C, D$  and  $E$  are the parameters of the regression model. Usually the value of  $E$  is set at 2 to enable calculation of the rest of the parameters by linear regression. Using the same vapor pressure data that were discussed in the previous section, the DIPPR staff obtained the following parameter values:  $A = 5.1857\text{E}+01$ ;  $B = -2.5987\text{E}+03$ ;  $C = -5.1283\text{E}+00$ ; and  $D = 1.4913\text{E}-05$ . In Figure 5, the residual plot of  $\ln(P_V)$  calculated by the Riedel equation using the DIPPR parameter values is shown. Observe that the residuals in this case are smaller by at least one order of magnitude than the residuals obtained using the Clausius-Clapeyron equation. There are two separate regions. In the first one (up to  $\ln(P_V) \sim 12$ ) the residuals are in the range of -0.04 to 0.02, not showing any particular trend. In the second region ( $\ln(P_V) > 12$ ) there is an order of magnitude reduction in the residual values (all smaller in absolute value than 0.003), however there is some trend in the residual values in this region. To obtain perfectly random distribution of the residuals, a different correlation should be used (the Wagner equation, for example). Improvement of the correlation by use of a different type of equation, however, is outside of the scope of the present paper.

The "copy and paste" facility is a very convenient and time-saving option when transferring numerical and other data between various data-bases and software packages. For example, the Riedel equation constants for various compounds can be copied into an Excel spreadsheet. After pasting the parameter values into Excel, they are stored with the exact same value as they appear in the DIPPR database, however they are displayed rounded to two three decimal digits:  $A = 5.19\text{E}+01$ ;  $B = -2.60\text{E}+03$ ;  $C = -5.13\text{E}+00$ ; and  $D = 1.49\text{E}-05$  when the default number display format is used. Copying these values from Excel and pasting it into a document or a different software package will results in a loss of the additional digits which were included in the DIPPR published values.

Calculating the vapor pressure using the rounded values of the parameters leads to significant errors in the calculated pressure values. The residual plots of  $P_V$  calculated by the Riedel equation using exact and rounded parameter values are shown in Figure 6. Observe that the residual (error) in the highest pressure range is of the order of 2.5% when rounded

parameter values are used in comparison to 0.3 % with exact parameter values. Thus, careless rounding of parameter values causes the prediction error, in this example, to increase by almost one order of magnitude.

## Conclusions

A complete definition of a problem should specify the parameter values and the uncertainty levels for these parameters. Specification of the uncertainty levels of the parameters enables a determination of the uncertainty of the solution. This provides justification to number of significant digits to be used in documentation of the results and sometimes (as in the example presented) it may reveal that the proposed process cannot be implemented in practice because of levels of uncertainty that are too high.

The correlation coefficient ( $R^2$ ) and a simple plot of calculated values versus experimental values can be misleading in determining the goodness of the fit between the data and the regression model. The residual plot is the most reliable means for examining the goodness of the fit.

A transfer of numerical data between various databases, documents and programs by electronic "copy and paste" may result in a loss of significant digits of important parameter values. Beware of unintended rounding during such transfer.

Proper attention to the potential causes of errors that are presented here and in previous publications<sup>[1],[2],[3],[4],[5]</sup> can considerably improve the reliability of the results in numerical problem solving and in the correlation of data by regression.

## References

1. Brauner, N., M. Shacham and M. B. Cutlip, "Computational Results: How Reliable Are They? - A Systematic Approach to Modal Validation", *Chem. Engr. Education*, **30**(1), 20-25 (1996).
2. Shacham, M., N. Brauner and M. Pozin, "Potential Pitfalls in Using General Purpose Software for Interactive Solution of Ordinary Differential Equations", *Acta Chimica Slovenica*, **42**(1), 119-124 (1995).
3. Shacham, M. and N. Brauner, "Correlation and Over-correlation of Heterogeneous Reaction Rate Data", *Chem. Engr. Education*, **29**(1) 22-25, 45 (1995).
4. Shacham, M., N. Brauner and M. B. Cutlip, "A Web-based Library for Testing Performance of Numerical Software for Solving Nonlinear Algebraic Equations", *Comput. Chem. Eng.* **26**(4-5), 547-554 (2002).
5. Shacham, M., N. Brauner, W. R. Ashurst, and M. B. Cutlip, "Can I Trust this Software Package? – An Exercise in Validation of Computational Results ", *Chem. Engr. Education*, **42**(1), 53-59 (2008).
6. Edgar, T. F., Himmelblau, D. M., and Lasdon, L. S., *Optimization of Chemical Processes*, 2nd Ed., McGraw-Hill, New York, 2001.
7. Froment, G.F., and Bishoff, K. B., *Chemical Reactor Analysis and Design*, 2<sup>nd</sup> Ed., Wiley, New York, 1990.
7. Parulekar, S. J., "Numerical Problem Solving Using Mathcad in Undergraduate Reaction Engineering", *Chem. Engr. Education*, **40** (1), 14 (2006).
8. Ingham, H.; Friend, D.G.; Ely, J.F.; "Thermophysical Properties of Ethane"; *J. Phys. Ref. Data*, **20**, 275(1991).

9. Rowley, R.L.; Wilding, W.V.; Oscarson, J.L.; Yang, Y.; Zundel, N.A. *DIPPR Data Compilation of Pure Chemical Properties Design Institute for Physical Properties*, (<http://dippr.byu.edu>), Brigham Young University Provo Utah, 2006.

Table 1. Isothermal CSTR solutions for various levels of uncertainty in the flow rate  $v_0$ .

	Feed flow rate ( $v_0$ ) L/s					
	0.2	0.2001	0.201	0.21	0.24	0.15
$C_A$ , 1st Root	<b>0.751536</b>	0.7533915 <sup>†</sup>	<b>0.7707207</b>	1.162332	-	0.341902
$C_A$ , 2nd Root	<b>2.130933</b>	2.124481	<b>2.066326</b>	1.311824	-	-
$C_A$ , 3rd Root	<b>8.117532</b>	8.122128	<b>8.162953</b>	<b>8.525843</b>	9.369219	-

<sup>†</sup>Correct digits are shown in **bold**

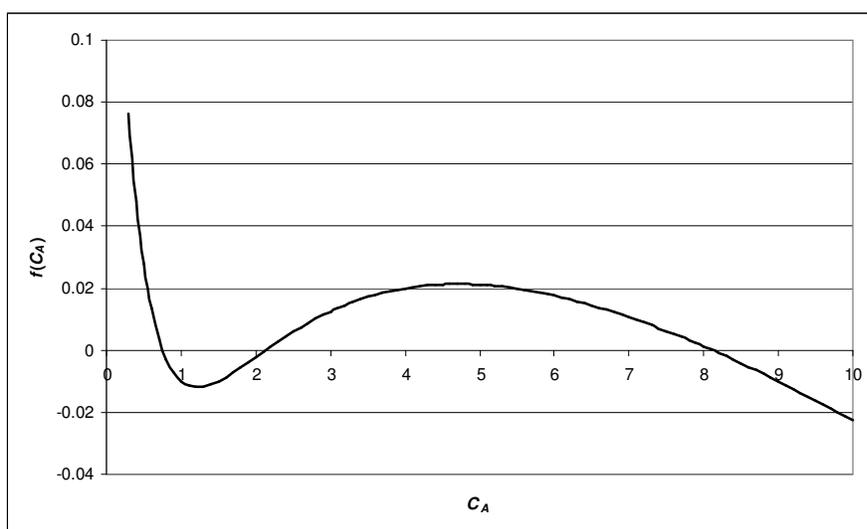


Figure 1. Plot of  $f(C_A)$  versus  $C_A$  for  $v_0 = 0.2$  L/s

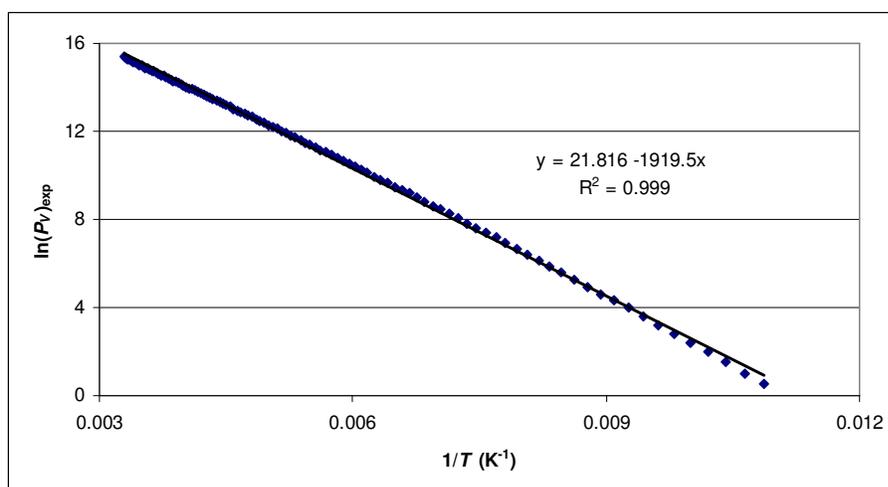


Figure 2. Plot of  $\ln(P_V)_{\text{exp}}$  versus  $1/T$  for ethane between the triple point and  $T_C$

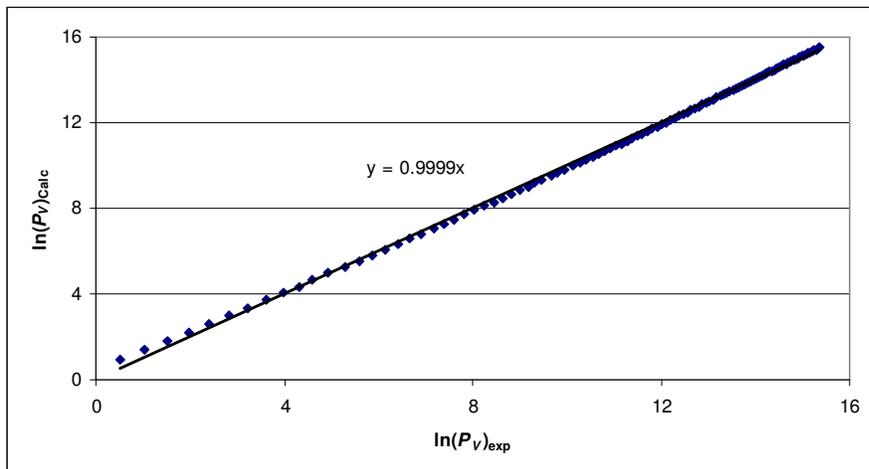


Figure 3. Plot of  $\ln(P_V)_{\text{calc}}$ , calculated by the Clausius-Clapeyron equation, versus  $\ln(P_V)_{\text{exp}}$  for ethane between the triple point and  $T_C$

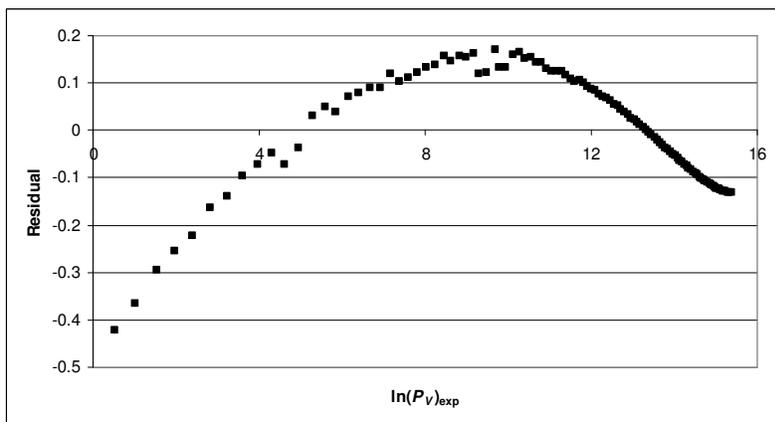


Figure 4. Residual plot for  $\ln(P_V)$  calculated by the Clausius - Clapeyron equation for ethane between the triple point  $T_C$

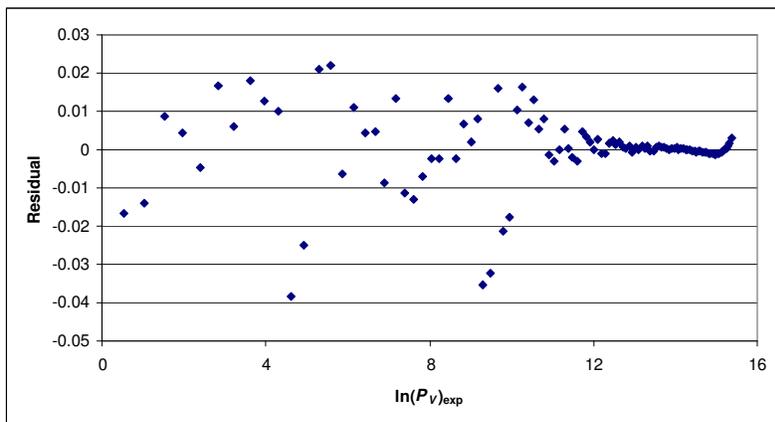


Figure 5. Residual plot for  $\ln(P_V)$  calculated by the Riedel equation for ethane between the triple point and  $T_C$

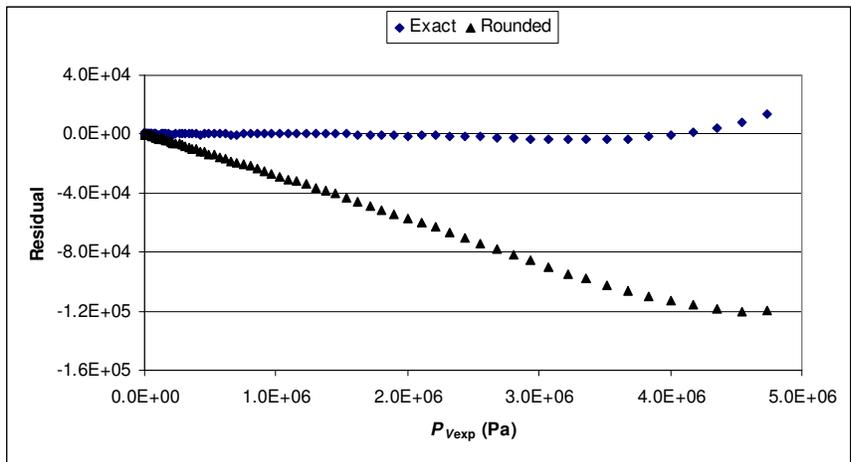


Figure 6. Residual plot for  $P_V$ , calculated by the Riedel equation, for ethane between the triple point and the  $T_C$