A First Course on Kinetics and Reaction Engineering
Unit 14. Differential Data Analysis

Overview

The design equations (reactor models) for the perfectly mixed batch reactor and for the PFR are differential equations. This presents a small problem when data from reactors of these types are used to test a potential rate expression. Experimental reactor data do not include values for the derivatives that appear in these equations, and as a result, the equations cannot be fit directly to the experimental data. Two approaches are commonly used to address this problem. This unit considers one of these two approaches wherein the experimental data are used to generate values for the derivatives, which then allows the differential design equations to be fit directly to the data. This is known as differential data analysis; its advantages, disadvantages and limitations are discussed herein.

Learning Objectives

Upon completion of this unit, you should be able to define, in words, the following terms:
- forward, backward and central differences
- initial rate

Upon completion of this unit, you should be able to write the defining equation for the following quantities:
- forward, backward and central differences

Upon completion of this unit, you should be able to perform the following specific tasks and be able to recognize when they are needed and apply them correctly in the course of a more complex analysis:
- Explain the difference between integral and differential data analysis for batch reactor and PFR data
- Describe the assumptions used in differential analysis of batch reactor and PFR data and its limitations
- Perform differential analysis of batch reactor and PFR data to test the validity of a rate expression and obtain the best values of the parameters appearing in it
- Describe how to measure initial rates with a batch reactor and perform differential analysis using initial rate data
- Describe the differential operation of a PFR
- Explain why differential data analysis can be used for any batch reactor data, but can only be used with PFR data if the reactor is operated differentially
- Distinguish between PFR data that can and cannot be analyzed using the differential method

Information

The differential method of data analysis first will be presented here with reference to the analysis of data from a batch reactor. Following that, application of the differential method of data analysis to PFR data will be considered. For present purposes it is assumed that experimental data have been collected using an isothermal, perfectly-mixed batch reactor. For illustrative purposes, it is further assumed that a

single reaction takes place at the conditions studied, and that the purpose of the experimentation is to find a suitable rate expression for that single reaction. Under these circumstances, the mole balance design equation for the reactor is given by equation (14.1). In equation (14.1), i denotes a reagent, j denotes a reaction, n denotes number of moles, V represents the volume within which the reaction occurs, t denotes time and $r_{i,j}$ is the rate of generation of species i via reaction j. Inherent in the use of this design equation is the assumption that the appropriate normalization factor for the reaction being studied is the volume.

$$\frac{dn_i}{dt} = Vr_{i,j} \tag{14.1}$$

The experimental data will include the full composition of the fluid at the start of the run. Beyond this, the data will consist of measured values some quantity related to the extent of reaction versus the elapsed time since the start of the run. These measured quantities are sufficient to calculate the moles, n_i, of any species, i, for each experimental data point.

Equation (14.1) is a differential equation, whereas curve fitting (least squares) usually requires an algebraic equation to be fit to the data. There are two approaches that can be taken in the fitting of equation (14.1) to the experimental data. These approaches are known as integral data analysis and differential data analysis. In integral data analysis, differential equation (14.1) is first solved, leading to an algebraic equation for n_i versus t, and that resulting equation is then fit to the experimental data. This unit examines the other approach: differential data analysis.

Generally, differential data analysis is less accurate than integral data analysis. Differential data analysis is useful when one has very high quality data (very accurate measurements with very little noise in the data) or when one simply wants to perform a "quick and dirty" preliminary analysis. The essence of differential data analysis lies in approximating a numerical value for the derivative, $\frac{dn_i}{dt}$, and then treating it like any other experimentally measured variable. When this is done, equation (14.1) (or a linearized form thereof) can be fit directly to the experimental data.

There are at least three ways to approximate $\frac{dn_i}{dt}$. The first is based upon the approximation given in equation (14.2). The approximation used in equation (14.2) is only valid when Δt is small; it becomes exact in the limit where Δn_i and Δt go to zero. For this reason equation (14.2) should only be used to approximate the value of $\frac{dn_i}{dt}$ when the changes in n_i and t between successive data points are small. The value of $\frac{dn_i}{dt}$ will normally be different for each data point. That is, $\frac{dn_i}{dt}$ will change continuously over the course of each experimental run.

$$\frac{dn_i}{dt} \approx \left. \frac{\Delta n_i}{\Delta t} \right|_{(t,n_i)} \tag{14.2}$$

The changes in the number of moles and the elapsed time, Δn_i and Δt, between each pair of data points can be computed directly from the (t,n_i) data points. If the value of $\frac{dn_i}{dt}$ was being estimated for data point k, one could use a forward difference, equation (14.3), a backward difference, equation (14.4)

or a central difference, equation (14.5). None of these finite differences approximations should be used if the data are "noisy" (the random fluctuations due to experimental uncertainty are relatively large). If the data are noisy, using these equations put too much emphasis on individual data points and miss the general trends of the data.

$$\left.\frac{dn_i}{dt}\right|_k \approx \frac{(n_i)_{k+1} - (n_i)_k}{(t)_{k+1} - (t)_k} \tag{14.3}$$

$$\left.\frac{dn_i}{dt}\right|_k \approx \frac{(n_i)_k - (n_i)_{k-1}}{(t)_k - (t)_{k-1}} \tag{14.4}$$

$$\left.\frac{dn_i}{dt}\right|_k \approx \left(\frac{1}{2}\right)\left(\frac{(n_i)_{k+1} - (n_i)_k}{(t)_{k+1} - (t)_k} + \frac{(n_i)_k - (n_i)_{k-1}}{(t)_k - (t)_{k-1}}\right) \tag{14.5}$$

A second way to approximate $\frac{dn_i}{dt}$ is graphical. One would first plot the data from the run as n_i vs. t, and draw a smooth curve through the data. Then, at the t corresponding to each experimental data point, a straight line tangent to the curve would be drawn. The slope of this tangent would be measured and taken as the approximate value of $\frac{dn_i}{dt}$ for that data point. The third way to approximate $\frac{dn_i}{dt}$ would be to fit a high order polynomial to the experimental data. Most spreadsheet programs have a built-in function that will do this. It is very easy to take the derivative of a polynomial, so this would be done, and the resulting equation would be evaluated at the t corresponding to each data point. The value so-computed would be taken as the approximate value of $\frac{dn_i}{dt}$ for that data point. The latter two methods of approximation, graphical and polynomial, are a little less susceptible to errors caused by noisy data than the finite differences approximation.

At this point, one will have a set of values for $\frac{dn_i}{dt}$ for each experimental data point k. The next step is to substitute the mathematical function that has been chosen for evaluation as a rate expression into equation (14.1). The resulting equation will contain one or more unknown parameters (rate constants, pre-exponential factors, activation energies, reaction orders, etc.). Fitting the equation (by linear or non-linear least squares) to the experimental data gives the "best" values for each of the unknown parameters, an estimate of the uncertainty in these values, and a correlation coefficient that indicates the "goodness" of the fit.

The final step is to make a decision whether the fit is sufficiently accurate. In making this decision one should not only consider the value of the correlation coefficient, but one should also use graphical analysis to look for systematic deviations between the experimental data points and the predictions of the mathematical function. If it is decided that the fit is sufficiently accurate, the data analysis is completed; otherwise, a different mathematical function is selected for analysis and the fitting is repeated until a good fit is obtained.

Initial rate data represent a special kind of differential data. Initial rate data are most easily generated when the reaction being studied involves two different reactants that can be prepared separately. In each experiment, one separately prepares the desired amount of each reactant at the desired reaction temperature so that the composition upon mixing the two will be known. The separate reactants are then rapidly mixed, and at some time very soon after, the reaction mixture is analyzed to determine the change in the amount of one of the reactants. The time at which the mixture was sampled and the measured change in the amount of one of the reactants is used to approximate the derivative as a forward difference (see above). Alternatively, several data points are measured in rapid succession. The moles versus time are calculated from these measurements and plotted. The slope of the plot at the origin, determined using any of the methods just described, is used as the initial rate. Because the value of the derivative is estimated at the initial time, the composition for the data point is taken to be the initial composition. The primary advantage of initial rate data is that the initial composition is more directly controlled and often, more easily varied.

To summarize, the differential analysis of batch reactor data involves the following steps:

- Approximate the value of $\dfrac{dn_i}{dt}$ for each experimental data point using one of the following methods
    - finite differences
    - graphical construction of tangents to the n_i vs. t data
    - fitting a polynomial to the n_i vs. t data, taking the derivative of the polynomial and evaluating it
- Substitute the mathematical function to be tested as a rate expression into equation (14.1)
    - The derivative is treated like any other experimentally measured variable
    - For a given data point, the composition used in the rate expression is the composition of the system at the point in time for which the derivative's value was estimated
- Fit the equation (14.1) to the rate-composition data and decide whether the fit is sufficiently accurate
- If the fit is sufficiently accurate
    - determine the best value for each unknown parameter
    - determine the uncertainty in the values of the parameters
- If the function is not sufficiently accurate, guess another function and repeat from that point on

The design equation for an isothermal, steady-state PFR with only one reaction taking place, equation (14.6), looks quite similar to the design equation for an isothermal batch reactor, equation (14.1). As such, it would seem that the differential method of data analysis for batch reactors can also be used with PFR data. In fact, PFR data can only be analyzed differentially if the reactor was operated differentially during the experimentation. The reason for this is related to the approximation of a numerical value for the derivative.

$$\frac{d\dot{n}_i}{dz} = \frac{\pi D^2}{4} r_{i,j} \qquad\qquad (14.6)$$

With the perfectly mixed batch reactor, the data that are collected experimentally are the same two variables that appear in the derivative, namely n_i and t. As a consequence, a plot of n_i versus t can be constructed from the data for any experiment. The slope of that plot can be measured at any time and

taken to approximate the value of $\dfrac{dn_i}{dt}$ at that time. The operating procedure for a PFR differs in that the experimentally measured data are not the same as the quantities that appear in the derivative of the design equation ($\dot{n}_i$ and z). Most commonly, all the experiments are done using one reactor, and so z has the same value in every experiment. That is, in those experiments, the composition is not typically measured at many points along the length of the reactor (that is, at different values of z), but instead it is only measured at the outlet from the reactor, $z = L$. Thus, one typically doesn't have the data for $\dot{n}_i$ versus z, and consequently a plot can't be generated and a slope can't be measured at any value of z. Simply put, the value of the derivative, $\dfrac{d\dot{n}_i}{dz}$, usually cannot be approximated directly from the experimental data.

When using a PFR, an approximate value for the derivative in equation (14.6) can only be obtained from the difference in composition between the inlet and the outlet as in equation (14.7), and then only under certain conditions. The approximation given in equation (14.7) is only valid mathematically when the difference between the inlet and outlet molar flow rates is small. *The practical consequence of this restriction is that the differential method of data analysis can only be used with PFR data if the PFR was operated so that the change in molar flow rate of species i was very small.* When plug flow reactors are operated in this way, it said that they are operated differentially. In order to use the differential method of data analysis, the reactor must have been operated differentially for every data point used in the analysis.

$$\frac{d\dot{n}_i}{dz} \approx \frac{\Delta \dot{n}_i}{\Delta z} = \frac{\dot{n}_i\big|_{outlet} - \dot{n}_i\big|_{inlet}}{L} \tag{14.7}$$

If the PFR is operated differentially during all experiments, then a value can be estimated for the derivative using equation (14.7), and the derivative in equation (14.6) can be treated as if it is a measured variable and not a derivative. That is, the data can be analyzed in the same way as differential analysis of batch reactor data, with one additional, minor difference.

When the differential method is applied to differential PFR data, the function being tested for its suitability as a rate expression is again substituted into the design equation. In this case, however, the derivative in the design equation is replaced by an average value for that particular experiment, equation (14.7). Consequently, the composition variables appearing in the rate expression should be their average values (i. e. between the inlet and the outlet) for that particular experiment. Put differently, for a given data point, the value of the derivative should be estimated using equation (14.7) and the corresponding composition should be taken to equal the average of the inlet and outlet composition. _The most important point, however, is that the reactor must have been operated differentially; otherwise the PFR data cannot be analyzed using the differential method._