Design Variable Selection to Simplify Process Calculations

WOOYOUNG LEE, JAMES H. CHRISTENSEN, and DALE F. RUDD

University of Wisconsin, Madison, Wisconsin

During the analysis of a large process system, the values of certain design variables must be preassigned to allow the solution of the design equations. A careless selection of these distinguished variables can lead to trouble, magnifying computational problems above what need be. A method is proposed for determining a set of design variables which could minimize the computational labor associated with process analysis. This work is part of a larger project on the computer aided design of process systems.

Process design is frequently burdened by a need for extensive computations; a burden which may not be inherent in the problem itself but which may result from our accidental approach to the problem. A careless first step may lead directly to horrendous computational problems which could easily be avoided. This paper deals with the strategy of making this first step, the selection of design variables. A simple algorithm is developed which enables the selection of a set of design variables which reduces the computational labor associated with the process analysis.

The process design equations can be expressed in symbolic form as

$$f_i(v_j) = 0 \qquad (1)$$
$$i = 1, 2, \ldots . N$$
$$j = 1, 2, \ldots . M$$

where $N < M$.

The design equations f_i tie together the process variables v_j through material and energy balances, equipment design equations, and the like. The equations constitute independent sources of information and may number into the hundreds for a typical chemical process. The variables (flow rates, equipment design parameters, temperatures, and the like) are even more numerous, the excess variables affording degrees of freedom for optimization.

The process analysis and optimization problem involves: the numerical specification of certain key design variables; the solution of the design equations for the unspecified state variables; the economic analysis of the results; and the respecification of the design variables to achieve more nearly the economic goals.

We wish to determine a set of design variables which simplifies the solution of the design equations. Our attention is now focused on the first two steps.

THE BIPARTITE GRAPH

We represent the system of equations by a special graph which exposes the structure of the equations compactly and lucidly, and which leads directly to their dissection. This useful notation is the bipartite graph of topology (*1, 5*).

Consider the following system of equations:

$$f_1(v_1, v_2, v_3) = 0$$
$$f_2(v_3, v_4, v_5) = 0 \qquad (2)$$
$$f_3(v_4, v_5, v_6) = 0$$
$$f_4(v_7, v_8, v_2) = 0$$

The corresponding graph is shown in Figure 1.

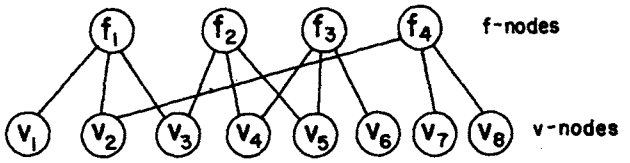


Fig. 1. The bipartite graph of Equations (2).

The equations form the f nodes, the variables form the v nodes, and an edge connects the equations to the corresponding variables.

The number of edges associated with a node defines ρ, the local degree of the node. The local degree of the f node f_1 is $\rho(f_1) = 3$ and the local degree of the v node v_5 is $\rho(v_5) = 2$.

This symbolic notation is the key to what follows; bipartite graphs have certain properties of extreme interest to the process engineer.

SELECTION OF DESIGN VARIABLES

Generally, it is necessary to specify the numerical values of certain of the variables in Equation (1) before solution can begin. The number of such distinguished variables is equal to the degrees of freedom of the system, that being the excess of variables over equations.

$$F = M - N \qquad (3)$$

Equation (3) is valid only when the design equations constitute independent sources of information, and it gives no hint as to which F variables should be chosen as design variables.

We are faced with a combinatorial problem, for there can be very many different ways to select F design variables from N total variables, not all of which lead to a soluble set of equations. A criterion is required for the selection of design variables.

Necessary Conditions

A necessary and sufficient condition for the solution of a set of equations is that the Jacobian determinant be nonzero (*2, 4*). While of considerable mathematical significance, this condition is of little direct use to the process engineer, since it is extremely difficult to compute the Jacobian determinant for a large system. We must look in directions other than toward the classical mathematics.

In a logically consistent set of N equations there will be N unspecified variables, and for each equation it will be possible to denote one variable as the output variable for that equation. And, no variable will be an output variable for more than one equation. If such an output assignment is possible, the equations are said to satisfy the diversity condition of Hall (3, 5), a necessary but not sufficient condition for the existence of a solution. This condition might be stated thus: *There must be at least* K *unspecified variables associated with each and every group of* K *equations, where* K = *1, 2, ...* N.

For example, the bipartite graph in Figure 2 describes an insoluble set of equations. Hall's condition fails at $K = 2$ where equations f_1 and f_2 have only v_1 associated with them.

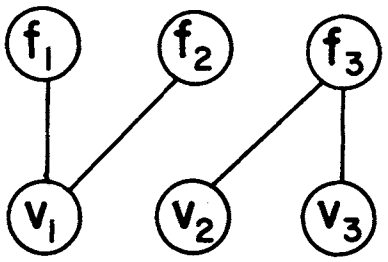
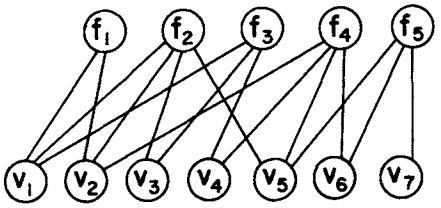


Fig. 2. The diversity condition not satisfied.

Armed with the diversity condition we consider the problem of selecting design variables which lead to a solvable set of equations. Later we consider the selection which leads to an easily solvable set of equations.

Assigning an output set. We now assign a direction to the edges of the bipartite graph, a step which identifies one of many possible orders of solution.

If a variable v_j *is* assigned as the output variable of equation f_i, then the edge connecting nodes f_i and v_j will be oriented from f_i to v_j and all other edges associated with f_i will be oriented to f_i.



F = 7 - 5 = 2

Fig. 3. The pipartite graph of Equation (3).

Design variables are not candidates for the output set, and their edges must be oriented to the equations. For example, consider the following equations and those in Figure 3.

$$\begin{aligned} f_1\,(v_1, v_2) &= 0 \\ f_2\,(v_1, v_2, v_3, v_5) &= 0 \\ f_3\,(v_1, v_3, v_4) &= 0 \\ f_4\,(v_2, v_4, v_5, v_6) &= 0 \\ f_5\,(v_5, v_6, v_7) &= 0 \end{aligned} \qquad (4)$$

If v_2 and v_3 are selected as the design variables, the subgraph in Figure 4 might result.

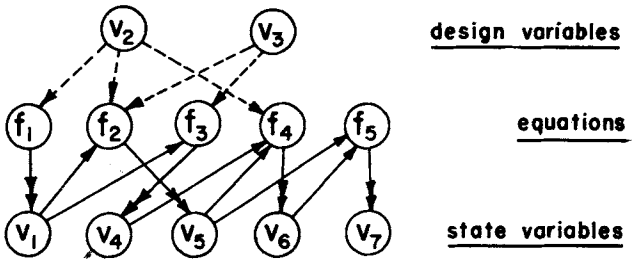


Fig. 4. V_2 and V_3 as design variables.

The fact that an output set has been assigned, denoted by the double-headed arrows, indicates that the diversity condition is satisfied. In addition, an order of solution is suggested by the direction of the arrows. The subgraph can be arranged as in Figure 5.

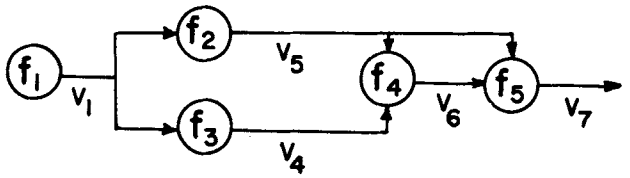


Fig. 5. The precedence order of equations.

This graph shows the precedence order of the system of equations (6). Starting from the left and preceding to the right, the set of five equations can be solved one at a time! There is no need for the simultaneous solution of equations, a troublesome task at best.

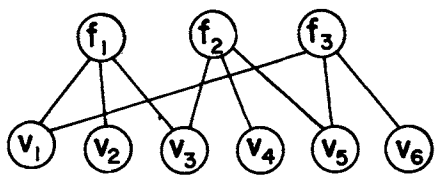
When such precedence order exists, the system does not possess a loop; the structure is said to be acyclic. Should a loop appear, the simultaneous solution of those equations within the loop must be undertaken (7).

The Effects of Design Variable Selection

The appearance of an acyclic structure in Figure 5 was not accidental; rather, it was the direct result of the selection of v_2 and v_3 as design variables. In this section we demonstrate the effects of design variable selection on the structure of the remaining computational problem.

For example, consider the following system and that in Figure 6.

$$\begin{aligned} f_1\,(v_1, v_2, v_3) &= 0 \\ f_2\,(v_3, v_4, v_5) &= 0 \\ f_3\,(v_5, v_6, v_1) &= 0 \end{aligned} \qquad (5)$$



F = 6 - 3 = 3

Fig. 6. The bipartite graph of Equations (4).

Case 1. Coupling Variables v_1, v_3, and v_5 are Design Variables: Once numerical values are assigned to these three variables the bipartite graph reduces to Figure 7. The equations can be solved one at a time, and the order of solution is immaterial.

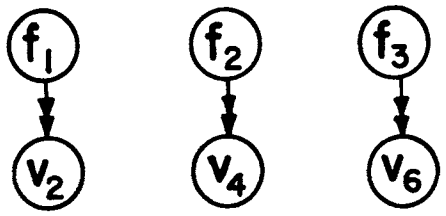


Fig. 7. V_1, V_3, and V_5 as design variables.

Case 2: Noncoupling Variables v_2, v_4, and v_6 are Design Variables: The bipartite graph becomes, for example, Figure 8. This unfortunate selection of design variables

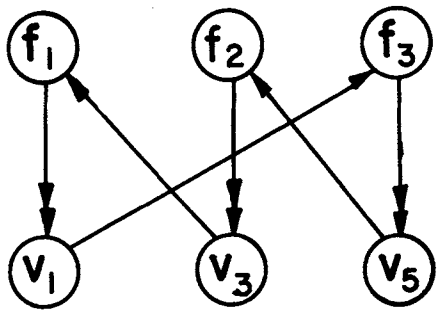


Fig. 8. V_2, V_4, and V_6 as design variables.

leads directly to a set of equations which must be solved simultaneously. We must now be troubled with the development of an appropriate iterative scheme of solution, with all the problems of convergence and the like. In the bipartite graph the order is *not* even unique, that is, the directions may be reversed.

Case 3: Variables v_1, v_4, and v_6 are Design Variables: The bipartite graph is now acyclic and the information flow is unique (Figure 9). Equation f_3 can be solved first for variable v_5, which then allows the solution of f_2 for variable v_3, which in turn enables the solution of equation f_1. The happy situation of case 1 has appeared again, in Figure 9.

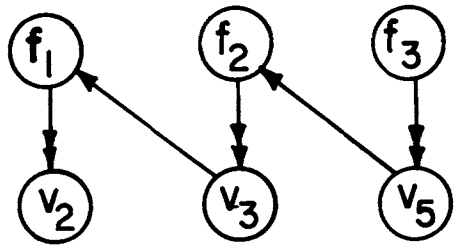


Fig. 9. V_1, V_4, and V_6 as design variables.

We temporarily assume that acyclic graphs are of the same difficulty equivalence class, and that cases 1 and 3 constitute equivalent selections of design variables. This is not strictly true, however, since certain equations can be solved more easily in one direction than the other. For example, if we have only a table of sin x, it is easier to solve $y = \sin x$ for y in terms of x, than x in terms of y. The same analogy holds true in process calculations. Later, this is accounted for by the assignment of difficulty scores in a report yet to be published.

SELECTION ALGORITHMS

It has been established that the careful selection of a set of design variables can alter the complexion of process design calculations. We now busy ourselves with altering the problem in our favor, toward a structure yielding an easily solved set of equations. A simple algorithm for design variable selection will be established.

A Selection Criterion

Generally, the labor of solving a set of design equations increases with the number of equations that must be solved simultaneously. We, therefore, define the following criterion for the selection of design variables:

A best set of design variables results in a structure in which the largest number of equations that must be solved simultaneously is minimized. A corollary of this criterion is that a best set of design variables results in an acyclic ordering of the design equations, if such an ordering is possible. Hence, if the difficulty of solving any equation is independent of the output variable and an acyclic assignment is possible, this criterion may give the minimum total effort for solving the set of design equations.*

An algorithm is now presented which selects a set of design variables, yielding an acyclic ordering of the design equations, if such an ordering is possible. The algorithm is based on the following observation:

The graph of an acyclic system must contain at least one v node v_j with $\rho(v_j) = 1$ and one f node f_i with $\rho(f_i) = 1$. This is so because every directed path in an acyclic graph must terminate, by definition, at a v node. Such a path cannot terminate at an f node, since each f node has one outgoing edge; therefore, it must terminate at a v node. Furthermore, since a v node can have only

* In a report yet to be published, we examine the validity of this criterion, and show that in situations where the difficulties of solving parts of the process design differ radically the concept of difficulty scores must be introduced.

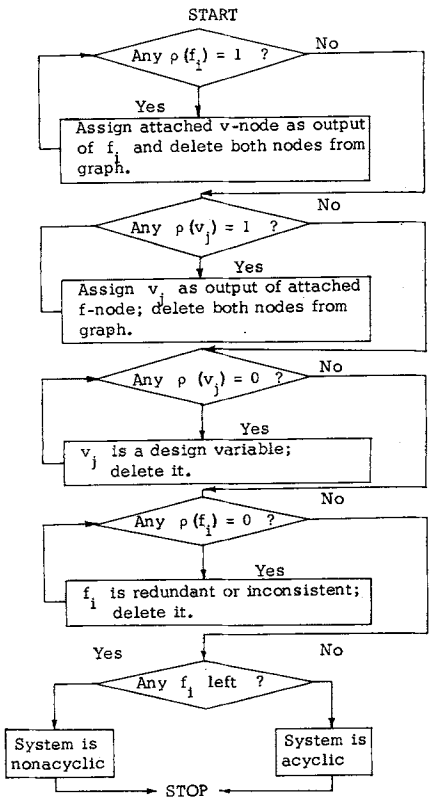


Fig. 10. Flow diagram of algorithm I.

one incoming edge, this v node must have exactly one edge, if a path terminates at this v node.

If this v node, and the f node of which it is an output, are deleted from the graph, the resulting subgraph must also be acyclic; hence, it will also contain at least one v node and one f node with one edge each.

Finally, we note that any f node connected to only one v node must necessarily have that v node as its output. We then have the algorithm for output and design variable selection shown in Figure 10.

Example of Algorithm I

Suppose that we have the system given in Figure 11. The design equations of the system are also given. We apply algorithm I to the graph of Figure 12. The system consists of a mixer, a reactor, and a separator. There are five streams in the system each described by two variables: the mass flow rate q and the composition c. The other variables in the system are the reactor volume V, the reaction temperature T, and the separation ratio S.

System:

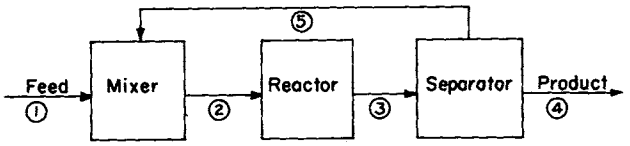


Fig. 11. A simple process system.

Design Equations:

Mixer:

$f_1\ q_1 + q_5 - q_2 = 0$	total mass balance
$f_2\ q_1c_1 + q_5c_5 - q_2c_2 = 0$	component balance

Reactor:

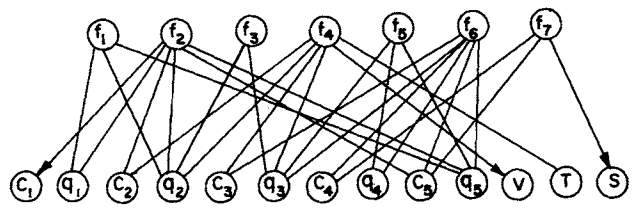
$f_3\ q_2 - q_3 = 0$	total mass balance
$f_4\ q_2c_2 - q_3c_3 - R(c_3, T, V) = 0$	component balance
	R is the reaction rate relationship

Separator:

$f_5\ q_3 - q_4 - q_5 = 0$	total mass balance
$f_6\ q_3c_3 - q_4c_4 - q_5c_5 = 0$	component balance
$f_7\ S = c_5/c_4$	separation ratio

The variables of interest are q_i's and c_i's, and V, T, and S. The number of design equations is $N = 7$. The number of variables is $M = 13$.

Graph:



F = 13 - 7 = 6

Fig. 12. The bipartite graph for the system.

Step 1: Look for nodes with the local degree of 1.

Step 2: Assign c_1, V, and S to f_2, f_4, and f_7, respectively.

Step 3: Eliminate these from the graph to get Figure 13.

Repeat steps 1 to 3: Figures 14 and 15 are obtained.

Design variables: T, c_2, c_4, c_5, q_3, and q_5

Order of calculations: See Figure 16.

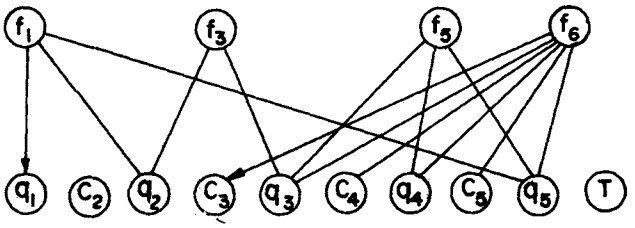


Fig. 13.

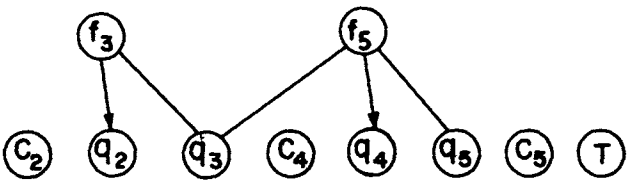


Fig. 14.

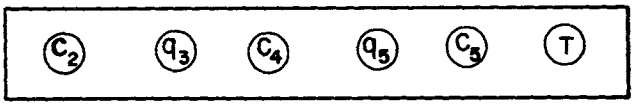


Fig. 15.

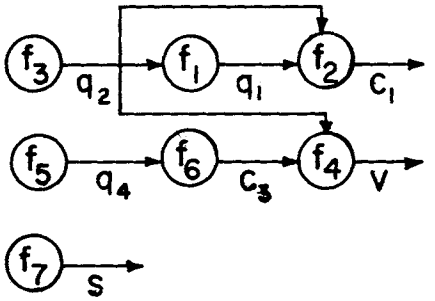
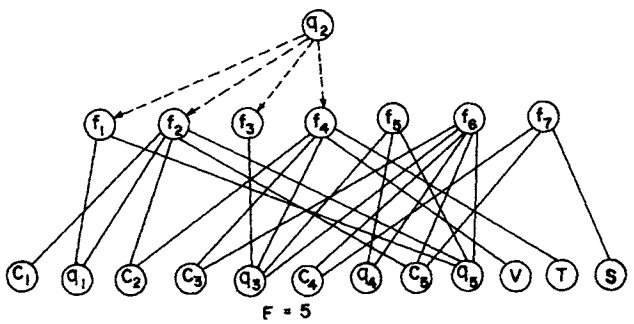


Fig. 16. The ordering of Figure 12.

Preferred Design Variables

For purely practical reasons the engineer may prefer to have close control over certain variables, and may elect to elevate such preferred variables to the level of a design variable regardless of the logical structure of the problem. For example, if quality control restrictions are imposed on a product concentration, the engineer may wish to compute only designs which meet those restrictions. This can be achieved by selecting that concentration as a design variable, the value of which is free to be selected directly.

This preference offers no difficulty. Such preferred variables are assigned prior to the use of algorithm I. If a restriction exists on the flow rate into the reactor, in the pre-



F = 5

Fig. 17. Preferred design variables.

vious example, q_2 might be a preferred variable. The selection of q_2 as a design variable consumes one degree of freedom and results in Figure 17, a graph more easily reducible by algorithm I.

Irreducible Systems

When algorithm I terminated, either an acyclic output assignment was produced, or $\rho(A) \geq 2$ for any node A in the remaining subgraph. If an acyclic assignment is possible, then the latter condition is impossible; this can be proved by contradiction. If the number of f nodes of this subgraph equals the number of v nodes, our degrees of freedom have been consumed and we are not free to select more design variables. A computational scheme for this subgraph has been already determined and the information flow can be traced out.

Should the degrees of freedom not be fully consumed, some more design variables are to be selected; however, an acyclic structure cannot be attained, regardless of our selection of the remaining design variables. The information flow can be changed acyclic only by tearing some of state variables. Suppose that if one assigns outputs to k f nodes and deletes these nodes and their outputs from the graph; then an acyclic assignment of the remaining subgraph may be possible by selecting appropriate design variables, where k is always less than the number of f nodes of the graph. It is said that k tears are supplied to the system and that the computation might be performed by a k dimensional iterative scheme. We wish to solve the system with the minimum number of tears k^*, since the computational burden of iterations increases with the dimension. Algorithm I cannot possibly find an acyclic ordering unless there is at least one v node with a single edge in the subgraph. This observation leads to the following algorithm.

Algorithm II

1. Apply algorithm I to the graph of the system.
2. If algorithm I succeeds in an acyclic ordering, then $k^* = 0$; otherwise, let $k = \min_{v_i} \rho(v_j) - 1$, for v_i in the remaining subgraph.
3. For each combination of k f nodes giving at least one v node with just one edge upon their deletion: (1) Apply algorithm I to the resulting subgraph; if no acyclic assignment is produced, go on to the next combination. (2) If an acyclic assignment is obtained, and k of the variables selected as design variables can be assigned as outputs of the deleted f nodes, then a minimum number of tears $k^* = k$ has been obtained; otherwise, go to the next combination.
4. If all combinations of k f nodes have been exhausted, increment k by one and go to step 3.

As an example, consider applying algorithm II to the simple system given in Figure 18. Obviously, algorithm I cannot reduce the size of the graph. We then have min $\rho(v_j) = 2$; hence, start with $k = 1$. Deletion of f_1 yields

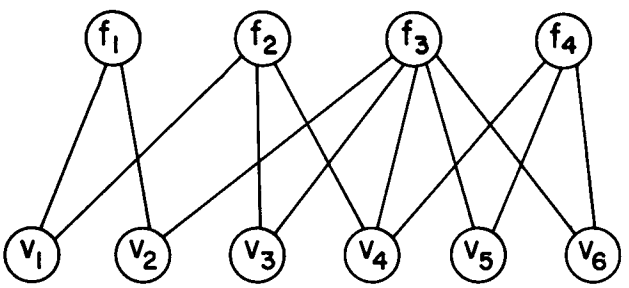


Fig. 18. Application of algorithm II.

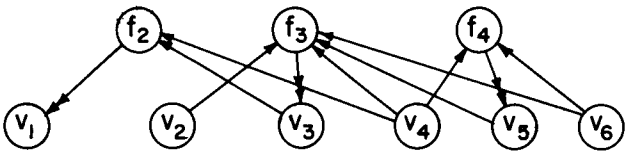


Fig. 19.

the following sequence of edge assignments by using algorithm I, shown in Figure 19.

Now v_2 can be assigned as the output of f_1 and $k^* = 1$. The system can then be arranged as in Figure 20 with design variables v_4 and v_6. In this case, a minimal tear was obtained at the first try; however, for large k^* the number of combinations to be considered may become excessively large. It is hoped that further investigation may yield some way to direct the search in order to reduce the number of cases which must be examined.

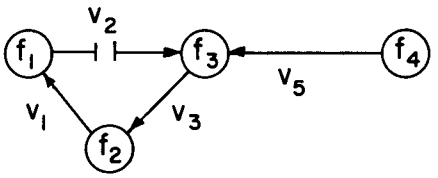


Fig. 20.

ORDERING THE PROCESS FLOW SHEET

Algorithms I and II were designed to aid the computer during the analysis of the design equations. We now extend these ideas to aid the engineer in the analysis of the process flow sheet, simplifying the process calculations. Only a slight extension is required to make this transition.

Figure 21 is the process flow sheet for an example process. The blocks represent unit operations, distillation, compression, filtration, heating, and so forth, and the arrows trace out the direction of flow of material and energy between the blocks. Notice that the direction of physical flow does not necessarily indicate the direction of information flow; the specification on the output product from stage M actually transfers information back into the system (8).Indeed, we take advantage of the nonuniqueness of the connecting arrows to reorder and simplify the system. In its present form, the specification of the values of the design variables leads directly to cumbersome iterative recycle calculations, a condition we shall avoid by the reassignment of new design variables.

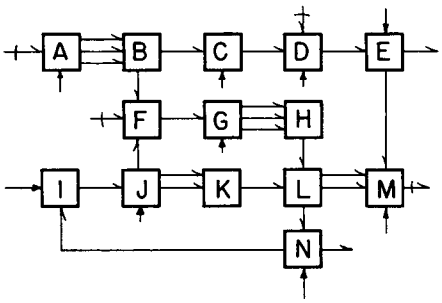


Fig. 21. A process flow diagram.

A block in the process flow sheet may be thought to embrace a number of design equations. For example, should the block represent a multicomponent distillation tower it may embrace the hundreds of equations required for an adequate description of that operation. To apply

the algorithms in their present form to such problems is unnecessary.

Notice that once the input variables to a block are specified the output variables can be solved for; this is the usual convention in flow sheet preparation. Thus, we may think of a block as a special equation which can possess several output variables, equal in number to the number of output arrows from the block in the flow sheet, but not necessarily those particular variables. With this interpretation we are in a position to apply the algorithms directly.

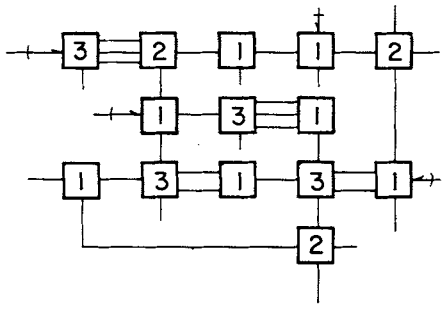


Fig. 22. An undirected flow diagram.

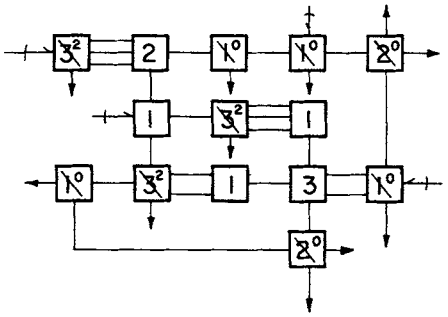


Fig. 23.

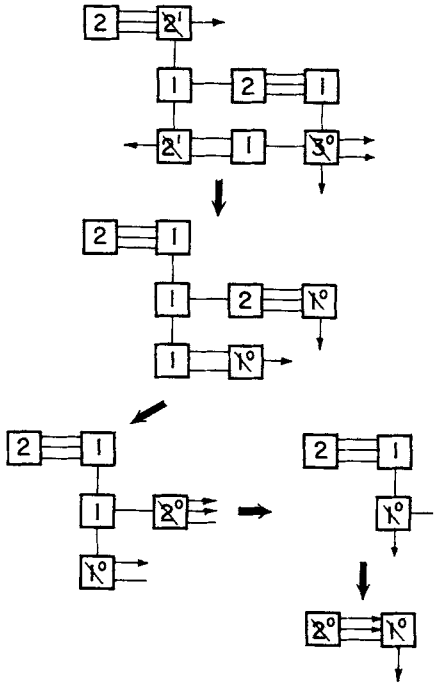


Fig. 24.

Figure 22 is the undirected counterpart of Figure 21 with the number of allowed outputs affixed. The variables which appear in only one block are denoted as outputs in Figure 23 with the allowed outputs reduced by the appropriate amount. Those blocks with no more allowed output are eliminated, and any new variables which now appear in only one block are then denoted as outputs from the reduced graph (Figure 24). When there are more single block variables than allowed outputs for a block, one has a choice which can only be made rationally by physical insight into the type of operation occurring in the block.

Finally, after all outputs have been assigned this information is placed on the original undirected graph. Any variables without assigned direction are design variables (Figure 25). Notice how this assignment has eliminated all recycle calculations and how the blocks now can be handled one at a time in the order prescribed in Figure 26.

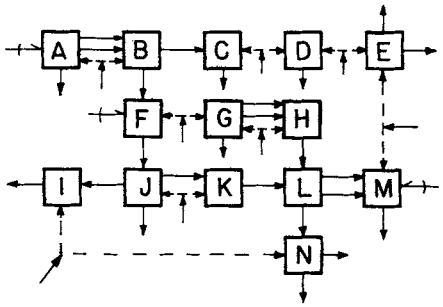


Fig. 25. A reassignment of design variables for Figure 21.

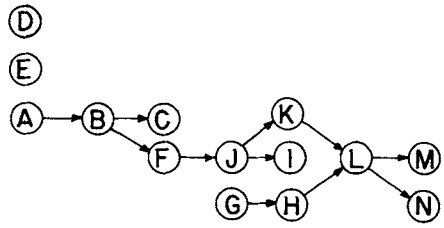


Fig. 26. A precedence ordering.

FINAL COMMENTS

We have demonstrated a systematic approach to the selection of design variables based on the criterion that acyclicity is desirable. System acyclicity will be obtained by use of these algorithms if it is obtainable merely by selecting the proper design variables. However, there are several points which require further comment.

It was implied in this report that information flow reversals can be made at any equation or any system component without difficulty. This is not true. It is well known that difficulties are often encountered should one attempt to *compute* certain variables, such as the kind of plates to use in a distillation tower, and that it is better to *specify* such variables and compute the consequences. The direction of information flow in the system determines the ease of component computation. Thus, in certain situations it may pay to tolerate cyclic structure in the system to ease the computation of a troublesome component.

However, before one can begin to understand the implications of partial acyclicity and the role that component *difficulty scores* may play in systems analysis, the limited case discussed in this paper must be presented. Hence, in this report the problem of systems analysis with completely

is measured only once. Any comparison to be made is not of multiple measurements of a single value but of one measurement relative to preceding and following measurements of some other value.

Whitaker and Pigford (*15*) showed with their method that, if one assumes: (1) that the fitting function used was the correct functional form of the data, (2) that the differences between the fitting function and the experimental data were due to random errors which follow a Gaussian distribution, and (3) that this distribution was the same for each point, then one can calculate a standard deviation of the first and second derivatives at every point. Then, using Student's t table, one may calculate the uncertainty in the derivatives corresponding to any required confidence level. They showed that for the example in their paper, the uncertainty in the first derivative at the 99% confidence level is $\pm$ 6%. If one continues their example to compute the second derivative, he sees that at the same confidence level the second derivative has an uncertainty of $\pm$ 200%.

Although this method of estimating the uncertainties of derivatives is useful as a caution against placing too much reliance on such a derivative, it is limited by the strong assumptions which are required to make such a calculation.

CONCLUSIONS

The problem of obtaining first and second derivatives from integral data is important in many rate process studies. Various methods of determining first and second derivatives from such data have been examined by means of three examples. Based on these examples, it appears:

1. Given a table of x-y data, the values of the derivatives computed from this table are a function of the method used to compute them. There is no method which automatically gives correct derivatives.
2. The first derivatives obtained by any of the methods are approximately equal. The second derivatives obtained by the various methods agree only if the data are very precise (that is, free from random or round-off error) and if the various methods fit the data within the limits of its precision.
3. If discontinuities exist in the second derivatives, the power series method will indicate this but will distort the entire second derivative in so doing. The other two methods will also indicate the discontinuity but without distorting the remaining part of the data.

THE PHILOSOPHY OF TREATING EXPERIMENTAL DATA

Experimental data like those in Table 1 are generally obtained at great effort and expense. As White and Churchill (*16*) have pointed out, the tendency of the experimenter is to use a data manipulation method which smooths over any experimental errors. The power series representation does this. On the other hand, the user of the data would like to know just how bad the data may be. The data difference method shows this.

Most workers prefer data interpretation schemes which are mechanical; that is, in which one feeds in the integral data and reads out the derivatives without further human intervention. The polynomial and segmented polynomial schemes fulfill this requirement. It is widely believed that these methods are objective and unbiased because the user makes no personal intervention in their application. However, in another sense, they are quite subjective and biased because, in applying them, the user has made strong assumptions (often without knowing it) as to the functional form of his data. If, as in example 3, there is a discontinuous second derivative and the polynomial method is used, the results will be very strongly influenced by the assumption of continuity of the second derivative.

In contrast, in the data difference method, the subjective, arbitrary parts of the procedure are all in plain view and must be made by the conscious intervention of the data interpreter. No a priori assumptions are made about the functional form of the data. The assumptions are made only after the data have been subjected to the mathematical manipulations which reveal its flaws.

ACKNOWLEDGMENT

The author would like to express his gratitude to Professor S. W. Churchill of the University of Michigan, who introduced the author to this subject and read and criticized the preliminary manuscript of this paper; to Dr. A. L. Tyler, who suggested the inclusion of the material on discontinuous derivatives; and to the University of Utah Computer Center, which donated computer time for the manifold calculations shown herein.

NOTATION

$a_o \ldots a_n$ = constants in power series fitting equation
$A_1 \ldots A_3$ = constants in Whitaker-Pigford fitting equations
P = pressure
r = residual = (observed value — calculated value)
t = time
V = volume per mole
x = distance or some unspecified coordinate
y = some unspecified coordinate
θ = angular distance

LITERATURE CITED

1. Churchill, S. W., and R. R. White. *Dechema-Monograph. Band*, **32**, 31-38 (1959).
2. Eggers, H. T. C., *Trans. Am. Inst. Chem. Engrs.*, **27**, 334-374 (1931).
3. Fimple, M. D., Program WRAP, Sandia Corp., Albuquerque, N. M.
4. Hildebrand, R. B., "Methods of Applied Mathematics," p. 227, Prentice-Hall Englewood Cliffs, N. J. (1952).
5. Lanczos, Cornelius, "Applied Analysis," p. 313, Prentice-Hall, Englewood Cliffs, N. J. (1956).
6. Lapidus, Leon, "Digital Computation for Chemical Engineers," p. 135, McGraw-Hill, New York (1962).
7. Michels, Antonius, and G. W. Nederbragt, *Physica*, **3**, 569-577 (1936).
8. Mickley, H. S., T. K. Sherwood, and C. E. Reed, "Applied Mathematics in Chemical Engineering," p. 25, McGraw-Hill, New York (1957).
9. *Ibid.*, p. 16.
10. Murphy, S. R., Ph.D. thesis, Univ. Utah, Salt Lake City (June, 1965). The data used in Example 1 are preliminary data from this project; they do not appear in the thesis, but the method of obtaining them does.
11. Perry, J. H., "Chemical Engineers' Handbook," 3 ed., p. 91, McGraw-Hill, New York (1950).
12. *Ibid.*, 4 ed., pp. 2-59, McGraw-Hill, New York (1963).
13. Running, T. R., "Graphical Calculus," George Wahr, Ann Arbor, Mich. (1937).
14. Tyler, A. L., Ph.D. thesis, Univ. Utah, Salt Lake City (June, 1965).
15. Whitaker, Stephen, and R. L. Pigford, *Ind. Eng. Chem.*, **52**, 185-187 (1960).
16. White, R. R., and S. W. Churchill, *A.I.Ch.E. J.*, **5**, 354-360 (1959).
17. Worthing, A. G., and Joseph Geffner, "Treatment of Experimental Data," p. 170, Wiley, New York (1943).
18. Wylie, C. R., Jr., "Advanced Engineering Mathematics," p. 130, McGraw-Hill, New York (1960).

Manuscript received April 5, 1965; revision received May 4, 1966; paper accepted May 5, 1966.

d^2y/dx^2. Thus, in an experiment involving such a second derivative, the experimental result is a strong function of the method used to calculate the second derivative from the x-y data, as will be illustrated. The common methods are:

1. Plot the data, draw a smooth curve through it, and read the slopes from this curve (*8*).
2. Fit the entire set of data with some fitting function and then differentiate this function.
3. Fit small segments of the data with some fitting function and then obtain local derivatives by differentiating the segmented functions (*15*).
4. Difference locally via a difference table, plot the differences, and smooth the differences, either graphically or numerically (*13*).

Each of these procedures involves smoothing processes, that is, the replacement of a measured value with another smoothed value. The derivatives are then based on the smoothed values. How these methods smooth data will be illustrated in several examples.

FIRST EXAMPLE

Smoothing the Data

In columns 1 and 2 of Table 1 are shown the observed angular positions of a particle moving outward in a centrifugal force field at equal time increments. The measurements were taken from an enlarged, multiple-exposure photograph of the particle trajectory (*10*).

The first step in analyzing any such data is to plot it as θ vs. t. On this plot (not shown), it appears that a straight line represents the data fairly well. One way of drawing that line is via the least squares fit. This has been done, and the resulting residuals are shown in column 3 of Table 1. The residuals form a consistent (nonrandom) pattern, suggesting that a better fit can be obtained with some curve rather than a straight line.

If method 1 is to be used, a french or ship's curve is selected and a smooth curve drawn through the data points. Unfortunately, there is no reproducible way of doing this; two engineers will draw different curves, and a single engineer will draw two different curves if he smooths the same set of data twice. Methods are available (*8*) for making the best estimate of the slope of such a curve once it is drawn, but no reproducible method is available for drawing the curve.

To use method 2, one normally selects a power series as his fitting function because of the convenient availability of computer programs for fitting such power series by least squares fit. The data in columns 1 and 2 of Table 1 were fit by such a program (*3*) by using five arbitrary constants; the residuals are shown in column 4 of Table 1. The constants for the fit are shown in Table 3.

To use method 3, some fitting procedure is needed. Examples are the Douglas-Avakian (*8*) and Whitaker-Pigford (*15*) methods. The Whitaker-Pigford method determines the first and second derivatives at a point by fitting a three-constant power series (a parabola) to five consecutive data points, with the point in question as the midpoint, by least squares. From the resulting power series constants, the local values of the first and second derivatives are computed. The calculated value of the fitting function at any one point has up to five different values, one for the fit where it is the midpoint, one for the fit where it is one point lower, etc. In column 5 of Table 1 are listed the residuals calculated from the averages of these representations by

$$r_i = \frac{\sum_{j=1}^{j=n} |r_{ij}|}{n} \qquad (1)$$

Here r_i is the average residual at point i, r_{ij} is the value of r_i calculated by using the j^{th} curve fit which includes point i, and n is the number of curve fits that include point i. N is 1 for the first and last points, 2 for the second and next to last, etc., up to a maximum of five for interior points.

Method 3 (the difference table method) may be construed as a variant of the third. The use of the difference table in numerical analysis is well known (*4, 6, 9, 18*), but its use to test and differentiate engineering data is ignored by most textbooks (*4, 6, 9, 18*), treated lightly by others (*11*), and argued against by some (*5, 12*). A method of data graduation, based on the difference table method shown here, has been proposed (*2, 13*). A difference table of the data in Table 1 is shown in Table 2.

The connection between methods 3 and 4 is that both fit short sections of the data by arbitrary functions. In the Whitaker-Pigford method, a parabola is fitted to five points. In the difference table method, the first divided differences ($\Delta\theta/\Delta t$ in Table 2) are obtained by fitting a straight line exactly to two consecutive points. The second divided differences ($\Delta^2\theta/\Delta t^2$ in Table 2) are obtained by

TABLE 1. DATA OF EXAMPLE 1

Observed and correlated angular positions of a spherical particle moving outward in a centrifugal force field (10).
Residual = θ Observed − θ Calculated.

(1) Observed angle θ, rad.	(2) Observed time t, sec.	(3) Residual from linear fit	(4) Residual from polynomial fit	(5) Residual from W-P fit	(6) Residual from data difference fit
0.251	0.0798	−0.0530	0.0027	0.0012	—
0.684	0.1596	−0.0162	−0.0050	0.0016	—
1.115	0.2394	0.0187	0.0002	0.0010	+0.0041
1.528	0.3192	0.0356	0.0012	0.0010	—
1.930	0.3990	0.0414	0.0042	0.0020	+0.0030
2.310	0.4788	0.0253	−0.0034	0.0027	—
2.692	0.5586	0.0112	0.0007	0.0016	+0.0046
3.060	0.6384	−0.0170	−0.0014	0.0006	—
3.427	0.7182	−0.0461	0.0009	0.0001	—
Avg. residual $= \sum_n \frac{\lvert r\rvert}{n}$		0.0293	0.0022	0.0013	0.0012
Root-mean-square residual $= \sqrt{\frac{\sum r^2}{n}}$		0.0369	0.0028	0.0015	0.0023

TABLE 2. DIFFERENCE TABLE FOR DATA SHOWN IN TABLE 1

θ, rad.	t, sec.	$\Delta\theta$, rad.	Δt, sec.	$\Delta\theta/\Delta t$, rad./sec.	$\Delta(\Delta\theta/\Delta t)$, rad./sec.	Δt, sec.	$\Delta^2\theta/\Delta t^2$, rad./sec.^{2}
0.251	0.0798						
		0.433	0.0798	5.42			
0.684	0.1596				−0.02	0.0798	−0.25
		0.431	0.0798	5.40			
1.115	0.2394				−0.22	0.0798	−2.75
		0.413	0.0798	5.18			
1.528	0.3192				−0.14	0.0798	−1.75
		0.402	0.0798	5.04			
1.930	0.3990				−0.28	0.0798	−3.50
		0.380	0.0798	4.76			
2.310	0.4788				+0.02	0.0798	+0.25
		0.382	0.0798	4.78			
2.692	0.5586				−0.12	0.0798	−1.50
		0.368	0.0798	4.62			
3.060	0.6384				−0.02	0.0798	−0.25
		0.367	0.0798	4.60			
3.427	0.7182						

fitting a parabola exactly to three consecutive points. The distinction between the two methods is that method 3 does the smoothing in this fitting procedure (three constants, five points), whereas method 4 uses the results of this exact fitting (three constants, three points) as the starting point for the smoothing process. To smooth the data by using the difference table, the divided differences $\Delta\theta/\Delta t$ shown in Table 2 are plotted vs. t in Figure 1. Each line represents the average value of $(d\theta/dt)_{obs.}$ over the t interval. The area under each line equals the change in θ for this Δt. Some of the properties of this type of plot have been published (*11, 13*).

In Figure 1 a straight line has been drawn through the various data lines. The lines scatter above and below it, suggesting experimental error. This is also obvious from the right-hand column of Table 2, in which the second divided differences are seen to vary erratically. For physical reasons it appears unlikely that the velocity of the particle is discontinuous, so it seems reasonable to smooth the data in Figure 1. Note that up to now no smoothing has been made at all in this method. In Figure 1 the value of $\Delta\theta/\Delta t$ from $t = 2$ to $t = 3$ is high; from $t = 3$ to $t = 4$, $\Delta\theta/\Delta t$ is slightly low relative to the rest of the data. If the measured value of θ at $t = 3$ were smaller, this would bring these more into line. Similar comments apply to the values of θ_5 and θ_7. A table similar to Table 2 is obtained by dropping these three readings completely out of the table and by repeating tabular differentiation. The resulting values of $\Delta\theta/\Delta t$ are plotted in Figure 2 and a smooth curve drawn through them. Thus, there are two smoothing steps in this method: (1) dropping the most unlikely data from the difference table and (2) drawing a smooth curve through the resulting first divided differences.

After drawing the smooth curve in Figure 2, one can now back-calculate the smoothed values of the data points which were dropped from this table. For example, the smoothed value of θ_3 must be

$$\theta_3 = \theta_2 + \left(\frac{d\theta}{dt}\right)_{\text{avg } t_2 - t_3} (t_3 - t_2) \qquad (2)$$

The average values of the derivatives are read from Figure 2 and the correction obtained. These are shown as residuals in column 6 of Table 2.

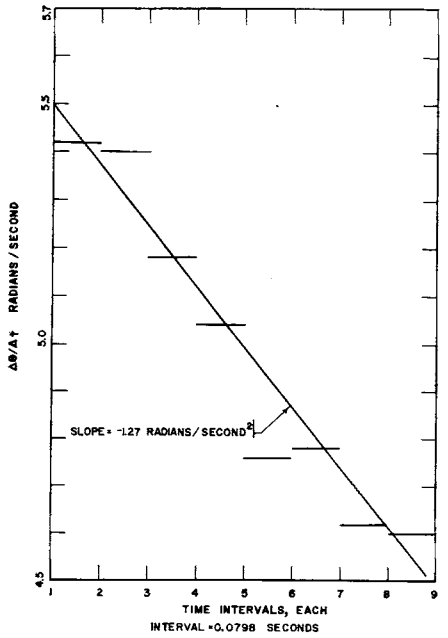


Fig. 1. $(\Delta\theta/\Delta t)$ vs. time from differenced data in Table 2.

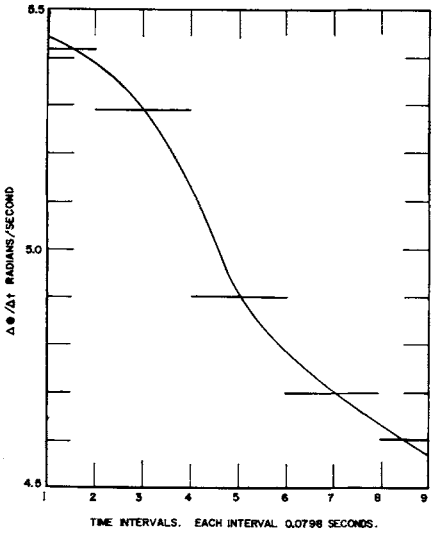


Fig. 2 $(\Delta\theta/\Delta t)$ vs. time from difference table made by dropping readings at t_3, t_5, and t_7 from Table 1.

Now the difference table is constructed again and the first and second derivatives are plotted to see if they form a smooth sequence. If not, a new curve is drawn in Figure 2 and the process repeated until smooth first and second derivatives are obtained.

Finding the Derivatives

After smoothing the data, as discussed, one can then find the derivatives from these smoothed data. If a smooth curve was drawn through the data (method 1), the slopes could be read by well-known methods (*8*). From the curve fit (method 2), by using a fourth-order power series, the derivatives are calculated from

$$\frac{d\theta}{dt} = a_1 + 2a_2t + 3a_3t^2 + 4a_4t^3 \qquad (3)$$

$$\frac{d^2\theta}{dt^2} = 2a_2 + 6a_3t + 12a_4t^2 \qquad (4)$$

Here a_n are the constants in the power series fit. These curves are shown in Figures 3 and 4.

From the segmented polynomial fit, by using the Whittaker-Pigford method (method 3), the local values of the first two derivatives at various points are calculated from equations of the form

$$\left(\frac{d\theta}{dt}\right)_j = A_{1j} + 2A_{2j}t \qquad (5)$$

$$\left(\frac{d^2\theta}{dt^2}\right)_j = 2A_{2j} \qquad (6)$$

Here A_{ij} are the i^{th} constants in the parabolic fit for the five points, including point j as the midpoint. These values are shown as point values in Figures 3 and 4.

From the difference table method (method 4), a new difference table is made of the smoothed data the resulting first and second derivatives shown in Figures 3 and 4. A smooth curve has been drawn through the data differences in Figures 3 and 4.

In Figure 3 it is clear that methods 2, 3, and 4 give about the same first derivative. Note that if the straight line through the data on the θ—t plot were accepted as the best representation of the data, then the values of $d\theta/dt$ corresponding to it would be a straight horizontal line $(d\theta/dt = 5.0)$ in Figure 3.

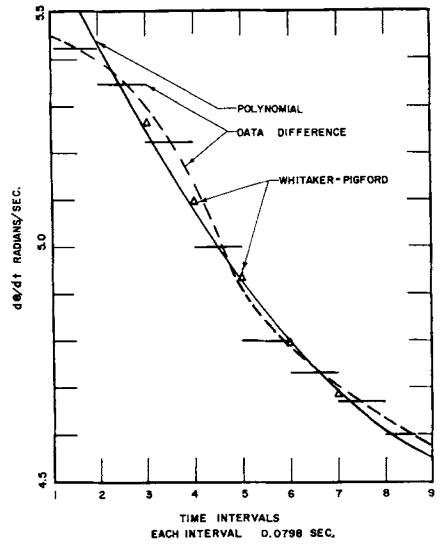


Fig. 3. Calculated first derivatives for first example.

In Figure 4, however, the second derivatives calculated by methods 2, 3, and 4 are quite different. Methods 3 and 4 give about the same derivatives, but the power series representation gives a much different derivative, particularly at low values of t. If the straight line on the θ—t plot were accepted as the best representation, then all values in Figure 4 would be zero.

Which Is Best?

In this case, the information actually sought was the second derivative, $d^2\theta/dt^2$. Figure 4 shows that by using different methods of smoothing the data to calculate this derivative from integral data, one can obtain results which differ significantly. Which should he believe, if any?

It seems clear that the straight-line representation on the θ—t plot is not the best. The average residual can be reduced tenfold by going from that representation to any of the others. Of the three remaining methods, the polynomial fit gives the largest average and root-mean-square residuals. Of all the methods, the polynomial method makes the most severe advance conditions on what form the data must have. If, as was done in this example, a fourth-degree polynomial is used, then, as shown in Equation (4), the second derivative, over the entire range of the data, is constrained to be a parabolic segment. The other two methods indicate that the second derivative does not have this shape.

On these two criteria, it appears that the polynomial fit is the worst of the three remaining methods. The other two give practically the same result; the choice between them should probably be based on convenience of application or on personal preference.

Do any of the methods give true derivatives? Given the data in Table 1 alone, there is no way to find true derivatives. All available methods smooth the data to find derivatives. Those which are most faithful to the data are those which change it least in smoothing and which make the fewest assumptions in advance about the form of the resulting smoothed data table. In smoothing these data, all methods indicate that changes larger than the experimental uncertainty are needed to produce a table with smooth derivatives. Thus, the experimental data must be significantly in error; only approximate derivatives are possible.

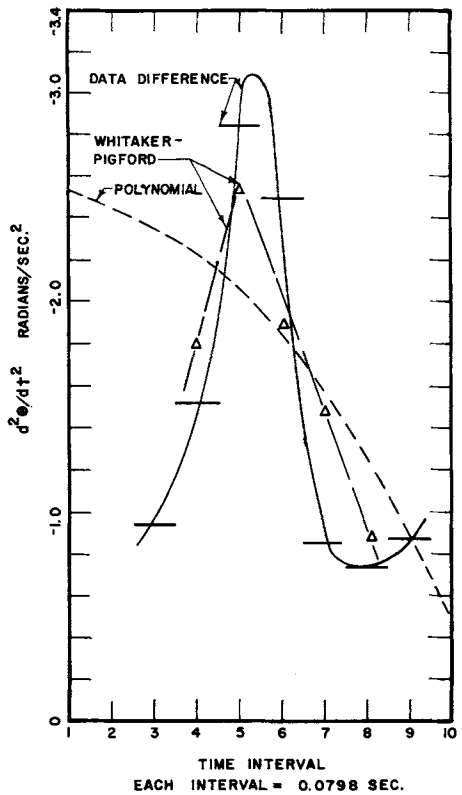


Fig. 4. Calculated second derivatives for first example.

SECOND EXAMPLE

The first example showed that for the data presented, four significant figures are not enough to provide an unambiguous second derivative. From Figure 4 it is also obvious that, starting with these data and using the different smoothing methods, one would calculate third derivatives which differ from each other in sign and in absolute value by several orders of magnitude.

If very precise data are smoothed by the various methods, no such ambiguity results. An example is shown in Figure 5. Here the P-V (pressure-volume) data of Michels and Nederbragt (7) for methane at 0°C. have been differenced three times and the values of $\Delta^3P/\Delta V^3$ plotted

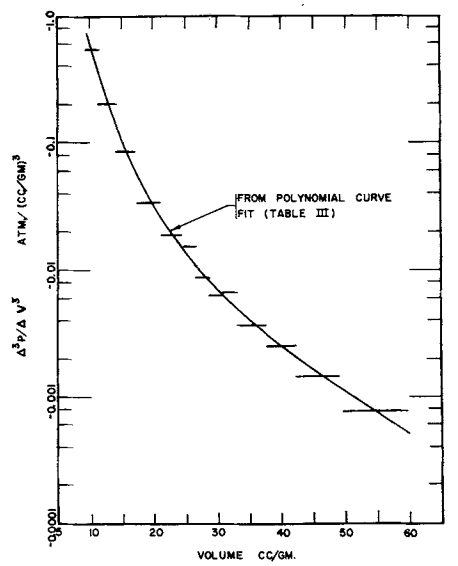


Fig. 5. $(\Delta^3P/\Delta V^3)$ of methane at 0°C. Data of Michels and Nederbragt (7).

vs. V. The results of random experimental errors (or round-off error) show up slightly in this plot. They are more apparent in the $\Delta^4P/\Delta V^4$ plot (not shown), but that plot also can be readily interpreted by drawing an average curve through the data differences.

Michels and Nederbragt (7) fitted their experimental data by a truncated power series. The constants are also shown in Table 3. By using this equation, d^3P/dV^3 was calculated, and is also shown in Figure 5. The agreement between the derivatives obtained by data differencing and by power series representation is very good. (The agreement is also satisfactory for the power series and data differenced values of d^4P/dV^4.) This example shows that if data of great precision are available, either method is acceptable (if the data can be accurately represented as a truncated power series). Data of this precision could also probably be plotted on huge graph paper and graphically differentiated with good accuracy. Notice that in order to obtain data precise enough to stand three differencings, it is necessary to use data from thermodynamics. Rate data of this precision are rare, if they exist at all.

Discontinuous Derivatives

Discontinuous first derivatives are common in thermodynamics (for example, the break in the cooling curve of a pure component, which corresponds to a phase change), but they are less common in rate processes. For example, a discontinuous velocity would require an infinite acceleration, which is physically unlikely. However, if measurements were made of temperature–distance in a double-pipe cooler-condenser in which a multicomponent mixture was being cooled and condensed, then the temperature–distance curve would be continuous but the dT/dx vs. distance curve would have a jump discontinuity at the point where condensation began.

Discontinuous second derivatives are more common in problems involving motion. For example, if one measured the distance–time behavior of an auto over the time span during which the driver took his foot off the accelerator and shoved it quickly onto the brake, the distance–time curve would be continuous, the velocity–time curve would be continuous, but the acceleration–time curve would have a jump discontinuity.

Customarily, mathematicians assume that all natural phenomena are representable by analytic functions. But analytic functions have continuous derivatives of all orders, whereas the previous examples show that significant physical phenomena have discontinuous derivatives. If one uses a derivative finding method, which assumes in advance that the data can be represented by an analytic function, then he has guaranteed in advance that, if such jump discontinuities in derivatives exist, they will be smoothed over. Tyler (14) has shown that the acceleration of spheres in a moving gas stream can be discontinuous. If he had used an analytic function method of differentiating his position–time data, he would never have observed this behavior.

TABLE 3. CONSTANTS IN DATA FITTING EQUATIONS

$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5$

Data set	Table 1 linear fit	Table 1 polynomial fit	Michel's data (7)
y	θ	θ	PV
x	t	t	ρ
a_0	−0.09211	−2.074934	1.00242
a_1	4.964	5.808533	-2.4144×10^{-3}
a_2	0	−1.208626	5.7541×10^{-6}
a_3	0	0.0	-3.550×10^{-9}
a_4	0	0.3208876	1.218×10^{-11}

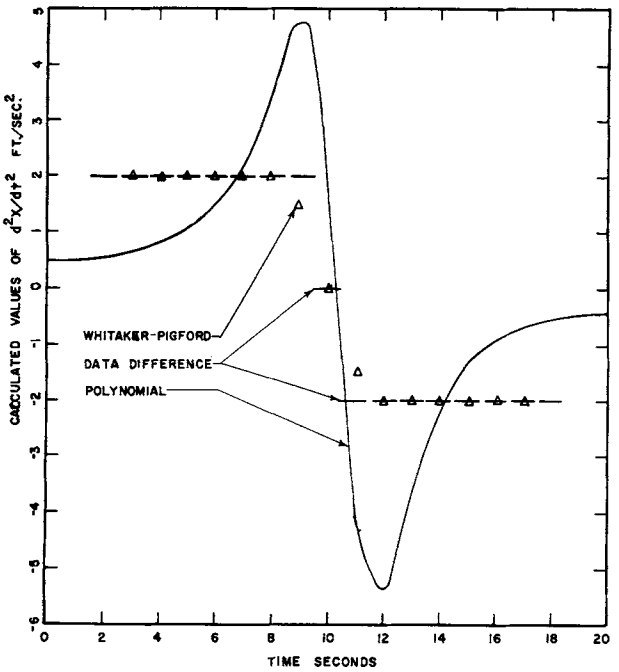


Fig. 6. Calculated second derivatives for discontinuous derivative example.

THIRD EXAMPLE

To illustrate the problem of obtaining second derivatives from integral data when the second derivative is discontinuous, a set of values was generated from the equations

$$\text{at } t = 0,\ x = 0, \text{ and } dx/dt = 0$$

$$\frac{d^2x}{dt^2} = 2\,\frac{\text{ft.}}{\text{sec.}} \text{ for } 0 \leq t \leq 10 \text{ sec.}$$

$$= -2\,\frac{\text{ft.}}{\text{sec.}} \text{ for } 10 < t \leq 20 \text{ sec.} \qquad (8)$$

The set of values gives x at twenty equal time intervals from 1 to 20 sec.

These values were then fit with a fifth-order polynomial of the form

$$t = a_o + a_1x \ldots + a_5x^5 \qquad (9)$$

(This gave a better fit, that is, lower residuals than $x = a_o + a_1t$, etc.) These values were also represented by the Whitaker-Pigford method. The second derivatives calculated by the polynomial, Whitaker-Pigford, and data difference methods are shown in Figure 6.

All three methods indicate a sharp change in the second derivative at $t = 10$. The data difference and Whitaker-Pigford methods do so without distorting the derivative in the rest of the data range. The polynomial fit, which asserts in advance that a jump discontinuity in the second derivative does not exist, distorts the derivative over the entire data range to make its assertion come true.

STATISTICAL APPROACHES

As shown above, all methods of calculating derivatives from integral data involve some smoothing before the calculation of derivatives. There are well-known statistical methods for deciding when one of a series of measurements can be rejected (17); it is pertinent to inquire if these can guide the above smoothing process. These statistical methods are based on comparing several measurements of the same value; they are not directly applicable to a table of x-y data, because in such a table each value

is measured only once. Any comparison to be made is not of multiple measurements of a single value but of one measurement relative to preceding and following measurements of some other value.

Whitaker and Pigford (*15*) showed with their method that, if one assumes: (1) that the fitting function used was the correct functional form of the data, (2) that the differences between the fitting function and the experimental data were due to random errors which follow a Gaussian distribution, and (3) that this distribution was the same for each point, then one can calculate a standard deviation of the first and second derivatives at every point. Then, using Student's t table, one may calculate the uncertainty in the derivatives corresponding to any required confidence level. They showed that for the example in their paper, the uncertainty in the first derivative at the 99% confidence level is $\pm$ 6%. If one continues their example to compute the second derivative, he sees that at the same confidence level the second derivative has an uncertainty of $\pm$ 200%.

Although this method of estimating the uncertainties of derivatives is useful as a caution against placing too much reliance on such a derivative, it is limited by the strong assumptions which are required to make such a calculation.

CONCLUSIONS

The problem of obtaining first and second derivatives from integral data is important in many rate process studies. Various methods of determining first and second derivatives from such data have been examined by means of three examples. Based on these examples, it appears:

1. Given a table of x-y data, the values of the derivatives computed from this table are a function of the method used to compute them. There is no method which automatically gives correct derivatives.
2. The first derivatives obtained by any of the methods are approximately equal. The second derivatives obtained by the various methods agree only if the data are very precise (that is, free from random or round-off error) and if the various methods fit the data within the limits of its precision.
3. If discontinuities exist in the second derivatives, the power series method will indicate this but will distort the entire second derivative in so doing. The other two methods will also indicate the discontinuity but without distorting the remaining part of the data.

THE PHILOSOPHY OF TREATING EXPERIMENTAL DATA

Experimental data like those in Table 1 are generally obtained at great effort and expense. As White and Churchill (*16*) have pointed out, the tendency of the experimenter is to use a data manipulation method which smooths over any experimental errors. The power series representation does this. On the other hand, the user of the data would like to know just how bad the data may be. The data difference method shows this.

Most workers prefer data interpretation schemes which are mechanical; that is, in which one feeds in the integral data and reads out the derivatives without further human intervention. The polynomial and segmented polynomial schemes fulfill this requirement. It is widely believed that these methods are objective and unbiased because the user makes no personal intervention in their application. However, in another sense, they are quite subjective and biased because, in applying them, the user has made strong assumptions (often without knowing it) as to the functional form of his data. If, as in example 3, there is a discontinuous second derivative and the polynomial method is used, the results will be very strongly influenced by the assumption of continuity of the second derivative.

In contrast, in the data difference method, the subjective, arbitrary parts of the procedure are all in plain view and must be made by the conscious intervention of the data interpreter. No a priori assumptions are made about the functional form of the data. The assumptions are made only after the data have been subjected to the mathematical manipulations which reveal its flaws.

ACKNOWLEDGMENT

The author would like to express his gratitude to Professor S. W. Churchill of the University of Michigan, who introduced the author to this subject and read and criticized the preliminary manuscript of this paper; to Dr. A. L. Tyler, who suggested the inclusion of the material on discontinuous derivatives; and to the University of Utah Computer Center, which donated computer time for the manifold calculations shown herein.

NOTATION

$a_o \ldots a_n$ = constants in power series fitting equation
$A_1 \ldots A_3$ = constants in Whitaker-Pigford fitting equations
P = pressure
r = residual = (observed value — calculated value)
t = time
V = volume per mole
x = distance or some unspecified coordinate
y = some unspecified coordinate
θ = angular distance

LITERATURE CITED

1. Churchill, S. W., and R. R. White. *Dechema-Monograph. Band*, **32**, 31-38 (1959).
2. Eggers, H. T. C., *Trans. Am. Inst. Chem. Engrs.*, **27**, 334-374 (1931).
3. Fimple, M. D., Program WRAP, Sandia Corp., Albuquerque, N. M.
4. Hildebrand, R. B., "Methods of Applied Mathematics," p. 227, Prentice-Hall Englewood Cliffs, N. J. (1952).
5. Lanczos, Cornelius, "Applied Analysis," p. 313, Prentice-Hall, Englewood Cliffs, N. J. (1956).
6. Lapidus, Leon, "Digital Computation for Chemical Engineers," p. 135, McGraw-Hill, New York (1962).
7. Michels, Antonius, and G. W. Nederbragt, *Physica*, **3**, 569-577 (1936).
8. Mickley, H. S., T. K. Sherwood, and C. E. Reed, "Applied Mathematics in Chemical Engineering," p. 25, McGraw-Hill, New York (1957).
9. *Ibid.*, p. 16.
10. Murphy, S. R., Ph.D. thesis, Univ. Utah, Salt Lake City (June, 1965). The data used in Example 1 are preliminary data from this project; they do not appear in the thesis, but the method of obtaining them does.
11. Perry, J. H., "Chemical Engineers' Handbook," 3 ed., p. 91, McGraw-Hill, New York (1950).
12. *Ibid.*, 4 ed., pp. 2-59, McGraw-Hill, New York (1963).
13. Running, T. R., "Graphical Calculus," George Wahr, Ann Arbor, Mich. (1937).
14. Tyler, A. L., Ph.D. thesis, Univ. Utah, Salt Lake City (June, 1965).
15. Whitaker, Stephen, and R. L. Pigford, *Ind. Eng. Chem.*, **52**, 185-187 (1960).
16. White, R. R., and S. W. Churchill, *A.I.Ch.E. J.*, **5**, 354-360 (1959).
17. Worthing, A. G., and Joseph Geffner, "Treatment of Experimental Data," p. 170, Wiley, New York (1943).
18. Wylie, C. R., Jr., "Advanced Engineering Mathematics," p. 130, McGraw-Hill, New York (1960).

Manuscript received April 5, 1965; revision received May 4, 1966; paper accepted May 5, 1966.